



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 109663

**TO: Karen A Lacourciere
Location: CM-1/11D09/11E12
Art Unit: 1635
Monday, December 08, 2003**

Case Serial Number: 09817913

**From: Toby Port
Location: Biotech-Chem Library
CM1-6A04
Phone: 308-3534**

toby.port@uspto.gov

Search Notes

Dear Examiner Lacourciere,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 13:02:22 ; Search time 585 Seconds

(without alignments)
11256.261 Million cell updates/sec

Title: US-09-817-913-2

Perfect score: 1611
Sequence: 1 atgtctggggtctctgccccg.....tccctcacgtttcttcccc 1611

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GeneBml:*

1: gb_ba:*

2: gb_ha:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_scs:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pin:*

35: em_hcg_rtd:*

36: em_hcg_mam:*

37: em_hcg_vrt:*

38: em_sy:*

39: em_hcg_hum:*

40: em_hcg_mus:*

41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1610.6	100.0	1611	6	AX703598
2	1596.6	99.1	1611	6	AX053101
3	1596.6	99.1	1611	9	HSU50079
4	1499.6	93.1	2093	9	BC000301
5	1498	93.0	2091	9	D50405
6	1498	93.0	2111	6	AR012007
7	1498	93.0	2111	6	E11455
8	1498	93.0	2111	6	E12388
9	1434.2	89.0	1449	12	BR007472
10	1240.2	77.0	1977	10	MMHIDE
11	1236.4	76.7	1999	10	MMU80780
12	1224.8	76.0	252509	2	AC098338
13	1201.8	74.6	253360	2	AC096943
14	1195.6	74.2	142347	10	AC125456
15	1195.6	74.2	181964	2	AC129023
16	1149	71.3	233552	2	AC131774
17	1149	71.3	238827	2	AC093467
18	1139.4	70.7	146907	2	AC141031
19	1139.4	70.7	223297	2	AC120669
20	1139.4	70.7	239100	2	AC105455
21	1139.4	70.7	249622	2	AC123200
22	1139.4	70.7	251643	2	AC125859
23	1137.2	70.6	214121	2	AC131629
24	1127.4	70.0	154474	2	AC117546
25	1112.8	69.1	182337	2	AC135881
26	1110.4	68.9	81975	2	AP000450
27	1110.4	68.9	146841	9	AC020549
28	1103.4	68.5	231632	2	AC119127
29	1098.4	68.2	187852	2	AC068041
30	1054	65.4	220688	2	AC134741
31	1054	65.4	235298	2	AC109842
32	1054	65.4	242577	2	AC121694
33	1054	65.4	280335	2	AC105705
34	1013.8	62.9	217804	2	AC117297
35	1013.8	62.9	248528	2	AC095666
36	997.4	61.9	1646	5	AF044169
37	997.4	61.9	1678	5	AF039751
38	992.6	61.6	1646	5	AF043328
39	984.2	61.1	257477	2	AC106951
40	972.8	60.4	160814	2	AC119545
41	972.8	60.4	260534	2	AC097994
42	953	59.2	1487	5	AF020658
43	953	59.2	2227	5	BC041296
44	943.4	58.6	2305	5	X1MB21
45	916.8	56.9	184614	2	AC114238

ALIGNMENTS:

RESULT 1	AX703598	1611 bp	DNA	linear	PAT 03-APR-2003
LOCUS	AX703598	Sequence 2 from Patent WO03006652.			
DEFINITION	AX703598				
ACCESSION	AX703598.1	GI:29538504			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Li, Z., Bonfils, C. and Beestman, J.				
TITLE	Inhibition of specific histone deacetylase isoforms				
JOURNAL	Patent: WO 03006652-A 2 23-JAN-2003;				

FEATURES
source
Methylgene, Inc. (CA)
Location/Qualifiers
1. 1611
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 428 a 389 c 439 g 354 t 1 others

ORIGIN

Query Match 100.0%; Score 1610.6; DB 6; Length 1611;
Best Local Similarity 100.0%; Pired. No. 0;
Matches 1611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 ATGCTGGGGGCTCTGCGCCGCTGCTGCTGCTCTCCCACTCGGTCACTCTGAGACACA 60
Db 1 ATGCTGGGGGCTCTGCGCCGCTGCTGCTGCTCTCCCACTCGGTCACTCTGAGACACA 60
Qy 61 GCCTGACGCRCTCTGTCACTCGGGGTAGACACCGGGGAGGCGAGCAAGATGCGGAGA 120
Db 61 GCCTGACGCRCTCTGTCACTCGGGGTAGACACCGGGGAGGCGAGCAAGATGCGGAGA 120
Qy 121 CGCAGGGGACCCCGAGGAAAGTCTGTTACTACTACGAGGGGAGATTGGAAATTAATT 180
Db 121 CGCAGGGGACCCCGAGGAAAGTCTGTTACTACTACGAGGGGAGATTGGAAATTAATT 180
Qy 181 ATGGACAAGGCGACCCCAATGAAAGCTTCAACGAAATCCGATGACTCATTAATTTGCTCTCA 240
Db 181 ATGGACAAGGCGACCCCAATGAAAGCTTCAACGAAATCCGATGACTCATTAATTTGCTCTCA 240
Qy 241 ACTATGCTCTCTACCGAATAATGAAATCTATCGCCCTCAACAAAGCCAAATGCTGAGAGA 300
Db 241 ACTATGCTCTCTACCGAATAATGAAATCTATCGCCCTCAACAAAGCCAAATGCTGAGAGA 300
Qy 301 TGACCAAGTACCAACGAGATGACTACTTAATTTCTTCCGCTCATTCGCTCCAGATACA 360
Db 301 TGACCAAGTACCAACGAGATGACTACTTAATTTCTTCCGCTCATTCGCTCCAGATACA 360
Qy 361 TGTGCGAGTACAGCAAGAGATGAGATTCAGGTTGAGAGAGCTGTCAGATTTTCG 420
Db 361 TGTGCGAGTACAGCAAGAGATGAGATTCAGGTTGAGAGAGCTGTCAGATTTTCG 420
Qy 421 ATGGCTGTGTGAGTTCTGTCACTGTTCTACTGTTGTTCTGTGGCAAGTCTGTGAAC 480
Db 421 ATGGCTGTGTGAGTTCTGTCACTGTTCTACTGTTGTTCTGTGGCAAGTCTGTGAAC 480
Qy 481 TTATTAAGCAGCAGACGACATCCCGGTGAATTTGGGCTGGGGGCTTGACCATTCAGAA 540
Db 481 TTATTAAGCAGCAGACGACATCCCGGTGAATTTGGGCTGGGGGCTTGACCATTCAGAA 540
Qy 541 AGTCGAGGGGATCTGGGCTTCTGTAGCAATGATATGTTGGCCATCTGGAACTGC 600
Db 541 AGTCGAGGGGATCTGGGCTTCTGTAGCAATGATATGTTGGCCATCTGGAACTGC 600
Qy 601 TAAAGTATCAACAGAGGCTGTGTACATTTGACATTTGATTTTCAATGTCGCGCTG 660
Db 601 TAAAGTATCAACAGAGGCTGTGTACATTTGACATTTGATTTTCAATGTCGCGCTG 660
Qy 661 AAGAGGCTTTTCAACACGAGCCGGGTCACTGCTGCTTTTCAATGATGAGAGT 720
Db 661 AAGAGGCTTTTCAACACGAGCCGGGTCACTGCTGCTTTTCAATGATGAGAGT 720
Qy 721 ACTTCCAGAGAACTGGGGACCTACGAGGATACCGGGGCTGGCAAGACAGATATTATGCTG 780
Db 721 ACTTCCAGAGAACTGGGGACCTACGAGGATACCGGGGCTGGCAAGACAGATATTATGCTG 780
Qy 781 TTAACATCCGCTCGAGACGGGATTTGATGACGATCTTATGAGCCATTTTCAAGCCG 840
Db 781 TTAACATCCGCTCGAGACGGGATTTGATGACGATCTTATGAGCCATTTTCAAGCCG 840
Qy 841 TCATGTCGAAAGTATGAGATGTTCCAGCCTAGTGGGCTGTTTCAAGTGGCTCAG 900
Db 841 TCATGTCGAAAGTATGAGATGTTCCAGCCTAGTGGGCTGTTTCAAGTGGCTCAG 900

```

Qy 901 ACTCCCTATCTGGGAGTCGGTATGTTAGTTCCTCAATCTATCTATCAAGACAGCCAAAGT 960
Db 901 ACTCCCTATCTGGGAGTCGGTATGTTAGTTCCTCAATCTATCTATCAAGACAGCCAAAGT 960
Qy 961 GTGTGGAATTTTCAAGAGCTTTTAACTGCTATGCTGATGCTGGGAGGCGGTGTTACA 1020
Db 961 GTGTGGAATTTTCAAGAGCTTTTAACTGCTATGCTGATGCTGGGAGGCGGTGTTACA 1020
Qy 1021 CCATTGTAACGTTGCCCCGGTGTGGACATATGAGACAGCTGTGGCCCTGATACGAGA 1080
Db 1021 CCATTGTAACGTTGCCCCGGTGTGGACATATGAGACAGCTGTGGCCCTGATACGAGA 1080
Qy 1081 TCCCTAATGAGCTTCCATACATGACTACTTTGAAATCTTTGAGACAGATTTTCAAGCTCC 1140
Db 1081 TCCCTAATGAGCTTCCATACATGACTACTTTGAAATCTTTGAGACAGATTTTCAAGCTCC 1140
Qy 1141 ACATGAGTCTTTCATATGACTATACCAAGACAGAAATGATACCTGGAGAAATCAAAAC 1200
Db 1141 ACATGAGTCTTTCATATGACTATACCAAGACAGAAATGATACCTGGAGAAATCAAAAC 1200
Qy 1201 AGGACTGTTTGAACCTTAGAATGCTGCGGACGACGACCTGGGGTCCAAACGACAGCGA 1260
Db 1201 AGGACTGTTTGAACCTTAGAATGCTGCGGACGACGACCTGGGGTCCAAACGACAGCGA 1260
Qy 1261 TTCTGAGAGCGCCATCCCTGAGAGAGTGGCGATGAGAGACGAAAGCAGCCTGACAAAGC 1320
Db 1261 TTCTGAGAGCGCCATCCCTGAGAGAGTGGCGATGAGAGACGAAAGCAGCCTGACAAAGC 1320
Qy 1321 GCATCTGATCTGCTCTCTGACAAACGAATTCCTGTGAGAAAGTTCTCCGATTTCTG 1380
Db 1321 GCATCTGATCTGCTCTCTGACAAACGAATTCCTGTGAGAAAGTTCTCCGATTTCTG 1380
Qy 1381 AAGAGAGGAGAGGGGGGCGGCAAGACCTTCAACTTCAAAAAAGCCAAAGAGTCA 1440
Db 1381 AAGAGAGGAGAGGGGGGCGGCAAGACCTTCAACTTCAAAAAAGCCAAAGAGTCA 1440
Qy 1441 AAACAGAGATGAAAAAGAAAGAACCCAGAGAGAAAGAAATCAACGAAAGAGAGA 1500
Db 1441 AAACAGAGATGAAAAAGAAAGAACCCAGAGAGAAAGAAATCAACGAAAGAGAGA 1500
Qy 1501 AAACAGAGAGAGAAAGCCAGAAAGCCAAAGGGGTCAAGAGAGGCGCAAGTTGGCTGAA 1560
Db 1501 AAACAGAGAGAGAGAAAGCCAGAAAGCCAAAGGGGTCAAGAGAGGCGCAAGTTGGCTGAA 1560
Qy 1561 TGAACCTTCCACTCTGCTGCTGCTGCTGAGTCCCTCAAGTTCTTTTCCCC 1611
Db 1561 TGAACCTTCCACTCTGCTGCTGCTGCTGAGTCCCTCAAGTTCTTTTCCCC 1611

RESULT 2
AX053101 1611 bp DNA linear PAT 12-JAN-2001
LOCUS
DEFINITION Sequence 25 from Patent WO0071703.
ACCESSION AX053101
VERSION AX053101.1 GI:12227157
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Macleod, A.R., Li, Z., and Besterman, J.M.
TITLE Inhibition of histone deacetylase
PATENT: WO 0071703-A 25 30-NOV-2000;
JOURNAL
Methylgene, Inc. (CA)
Location/Qualifiers
1. 1611
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 428 a 385 c 440 g 358 t
ORIGIN

Query Match		99.1%	Score 1596.6	DB 6	Length 1611	
Best Local Similarity		99.4%	Pred. No. 0			
Matches 1602		Conservative	0	Mismatches	9	Indels
					0	Gaps
						0
Qy	1	ATGTCGAGGGCTCTCTGCGCGCTGGTGGTGTGCTGTCTCCACATCGGTCAATCTGTGAAACACA	60			
Db	1	ATGTCGTGGGGTCTCTGCGCGCTGGTGGTGTGCTGTCTCCACATCGGTCAATCTGTGAAACACA	60			
Qy	61	GCCGTGACGRCCTCTGTCACTTCGGGGTGAACAACGCGGGAGGCGAGACAGATGGCGCAGA	120			
Db	61	GCCGTGACGRCCTCTGTCACTTCGGGGTGAACAACGCGGGAGGCGAGACAGATGGCGCAGA	120			
Qy	121	CGCAGGGCACCCGGAGGAAAGTCTGTATCTATCTACGACGGGGATGTTGGAATTAATTAATT	180			
Db	121	CGCAGGGCACCCGGAGGAAAGTCTGTATCTATCTACGACGGGGATGTTGGAATTAATTAATT	180			
Qy	181	ATGACAAAGGCAACCCATGAAAGCTCACGGAATCCCATATCATATATTCGTGCTCA	240			
Db	181	ATGACAAAGGCAACCCATGAAAGCTCACGGAATCCCATATCATATATTCGTGCTCA	240			
Qy	241	ACTATGGTCTCTACCGAAAAATGGAATCTATCGCCCTCACAAAGCCATGCTGAGAGA	300			
Db	241	ACTATGGTCTCTACCGAAAAATGGAATCTATCGCCCTCACAAAGCCATGCTGAGAGA	300			
Qy	301	TGACCAAGTACCAAGGATGACATCAATTAATCTTGCGCTCATCCGTCCAGATTAACA	360			
Db	301	TGACCAAGTACCAAGGATGACATCAATTAATCTTGCGCTCATCCGTCCAGATTAACA	360			
Qy	361	TGTGCGAGTACAGACAGACAGATGACAGATTCAACGTTGGTGAAGGACTGTCCAGTATTCG	420			
Db	361	TGTGCGAGTACAGACAGACAGATGACAGATTCAACGTTGGTGAAGGACTGTCCAGTATTCG	420			
Qy	421	ATGGCTGTGTTGAGTTCGTCAAGTGTCTACTGATGGTCTGTGGCAAGTCTGTGAAAC	480			
Db	421	ATGGCTGTGTTGAGTTCGTCAAGTGTCTACTGATGGTCTGTGGCAAGTCTGTGAAAC	480			
Qy	481	TTAATTAAGCAGACAGCGCACTGCCCGTGAATTTGGGCTGGGGGCTGTACCATGCAAGA	540			
Db	481	TTAATTAAGCAGACAGCGCACTGCCCGTGAATTTGGGCTGGGGGCTGTACCATGCAAGA	540			
Qy	541	AGTCCGAGGCACTGGGCTTCGTGTAACGTAATGATAGTCTTGAGCCATCTCGGAACTGC	600			
Db	541	AGTCCGAGGCACTGGGCTTCGTGTAACGTAATGATAGTCTTGAGCCATCTCGGAACTGC	600			
Qy	601	TAAAGTATCAACAGAGGGTGTCTGTACATTTGATATTCACCATGATGACGCGCTGG	660			
Db	601	TAAAGTATCAACAGAGGGTGTCTGTACATTTGATATTCACCATGATGACGCGCTGG	660			
Qy	661	AAGAGGCTTCTACACACGCGAGCGGGCTATGACTGTCTCTTCAATAGTATGAGAGT	720			
Db	661	AAGAGGCTTCTACACACGCGAGCGGGCTATGACTGTCTCTTCAATAGTATGAGAGT	720			
Qy	721	ACTTCCACAGAACTGGGGGCTACGGGGATACGGGGGCTGGGAAAGCAAGTATATGCTG	780			
Db	721	ACTTCCACAGAACTGGGGGCTACGGGGATACGGGGGCTGGGAAAGCAAGTATATGCTG	780			
Qy	781	TTAATCAACCGCTCCGAGACGGGATTTGATGACGAGTCTCATGAGGCCATTTTCAACGCGG	840			
Db	781	TTAATCAACCGCTCCGAGACGGGATTTGATGACGAGTCTCATGAGGCCATTTTCAACGCGG	840			
Qy	841	TCATGTCCAAAGTATGAGATGTTTCCAGCTTATGTCGGTGTCTTACATGATGTGCTCAG	900			
Db	841	TCATGTCCAAAGTATGAGATGTTTCCAGCTTATGTCGGTGTCTTACATGATGTGCTCAG	900			
Qy	901	ACTGCCATCTGGGGATCGTTTGGTTCCTCAATCTATCAATCAAGACACGCAAGT	960			
Db	901	ACTGCCATCTGGGGATCGTTTGGTTCCTCAATCTATCAATCAAGACACGCAAGT	960			
Qy	961	GTTGGAATTTGTCAAGACTTTAACTTGCCATCTGATGCTGGGAGCGCGTGGTTTACA	1020			
Db	961	GTTGGAATTTGTCAAGACTTTAACTTGCCATCTGATGCTGGGAGCGCGTGGTTTACA	1020			
Qy	1021	COATTCGTACGTTGCCCGGTGCTGGAATATAGACAGACTGTGGCCCTCGATACGAGAGA	1080			

Db	1021	CCATTGCTAAGCTTTGCCGGTGGTGGACATATGACAGCTGTGGCCCTGATACGGAGA	1080
QY	1081	TCCCTAATGAGCTTCCATACAAATGACTACTTTGAATTCCTTGGACCAAGATTTCAAGCTCC	1140
Db	1081	TCCCTAATGAGCTTCCATACAAATGACTACTTTGAATTCCTTGGACCAAGATTTCAAGCTCC	1140
QY	1141	ACATCAGTCTCTTCCAAATATGACTTACCAAGAACCAAGATGATCTGTGAGAAAGATCAAC	1200
Db	1141	ACATCAGTCTCTTCCAAATATGACTTACCAAGAACCAAGATGATCTGTGAGAAAGATCAAC	1200
QY	1201	AGGACCTGTTTGGAAACCTTAGAATGTGCGGACGCACTGTGGGATCCAAACGAGGCGA	1260
Db	1201	AGGACCTGTTTGGAAACCTTAGAATGTGCGGACGCACTGTGGGATCCAAACGAGGCGA	1260
QY	1261	TTTCCTGAGACGCGCATCCCTGAGAGAGTGGCGATGAGAGCGAAGACGACCTCTGACAAAG	1320
Db	1261	TTTCCTGAGACGCGCATCCCTGAGAGAGTGGCGATGAGAGCGAAGACGACCTCTGACAAAG	1320
QY	1321	GCAATCTCGATCTGTCTCTCTGACAAACGAAATTTGCTGTGAGAAAGATTTCTCGAATTTCTG	1380
Db	1321	GCAATCTCGATCTGTCTCTCTGACAAACGAAATTTGCTGTGAGAAAGATTTCTCGAATTTCTG	1380
QY	1381	AAGAGAGGAGAGAGGGGGGCGGCAAGAACTTTTCCAACTTTCAAAAAAAGCCAGAGACTCA	1440
Db	1381	AAGAGAGGAGAGAGGGGGGCGGCAAGAACTTTTCCAACTTTCAAAAAAAGCCAGAGACTCA	1440
QY	1441	AAACACAGGATGAAAAAAGAAAAAGATCCACAGAGGAAAGAAAGAAATACCCGAGAGAGGA	1500
Db	1441	AAACACAGGATGAAAAAAGAAAAAGATCCACAGAGGAAAGAAAGAAATACCCGAGAGAGGA	1500
QY	1501	AAACCAAGAGAGAGAACCCGAAGCCCAAAGGGGTCAAGAGAGAGGCCAAGTTGGCCCTGAA	1560
Db	1501	AAACCAAGAGAGAGAACCCGAAGCCCAAAGGGGTCAAGAGAGAGGCCAAGTTGGCCCTGAA	1560
QY	1561	TGAGCCTCTCAGCTGTGGCTTCTCTGTGAGTCCCTCAGCGTTTCTTTTCCCC	1611
Db	1561	TGAGCCTCTCAGCTGTGGCTTCTCTGTGAGTCCCTCAGCGTTTCTTTTCCCC	1611

```

RESULT 3
HSU50079          1611 bp      mRNA      linear      PRI 16-MAY-1996
LOCUS
DEFINITION      Human histone deacetylase HD1 mRNA, complete cds.
ACCESSION      U50079
VERSION      U50079.1
KEYWORDS      GI:1277083
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Taunton,J., Hassisg,C.A. and Schreiber,S.L.
TITLE      A mammalian histone deacetylase related to the yeast
transcriptional regulator Rpd3p
Science 272 (5260), 408-411 (1996)
JOURNAL
MEDLINE      96185439
PUBMED      8602529
REFERENCE
AUTHORS      Taunton,J., Hassisg,C.A. and Schreiber,S.L.
TITLE      Direct Submission
Submitted (27-FEB-1996) Jack Taunton, Chemistry, Harvard
University, 12 Oxford, Cambridge, MA 02138, usa
JOURNAL
FEATURES
Source
1..1611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="Jurkat T cell"
111..1559
/note="similar to S. cerevisiae RPD3, a global
transcriptional regulator, texpoxin receptor"
/codon_start=1

```

```

/product="histone deacetylase HD1"
/protein_id="AAC50475.1"
/db_xref="GI:1277084"
/translation="MAOTGRKRVKVVYDGDVNVYVGGHPPKPRIRMTNLLIN
YGLRKMELIRPRKABEMTKHSDYIIFLNSIRPDNNSESKOMQRNVGDCV
FDGLFEFCQLSTGGSVASAVLNKQDTDIANVAGDLHAKKSASFVYNDIULAI
LELLKQVRYLVIDIDIHGDVEAEAYTDRVMTVFPHKVEFPGTGLRDIQAK
GKYAVNVPLRDLIDDESVEAI FKPVMSKVMEMFQPSAVLQCSDSLSDPLGCFNL
TIKHAQCEVFKSFNLPMMLGGGCTIRVARCWYETVALDTEI PHELPNDYF
EYRPPDKHISPSNMNTNONTVEYLEKIKORLPENLPMLEPHAPVOMQAL PEALIPBE
SGDEDDPDKRISICSDKRIACEEFSBEEBEGGRKNISNFKAKVKTEDEKE
KDPEKKEVTEETKEKEPEAKGVKEEVKLA"
BASE COUNT      428 a      385 c      440 g      358 t
ORIGIN
Query Match      99.1%; Score 1596.6; DB 9; Length 1611;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1602; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGTCGGGGCTCTGCGCCGCTGGTGTCTCTCCCACTCGGTCACTCTGAGAACCA 60
DB 1 ATGTCGGGGCTCTGCGCCGCTGGTGTCTCTCCCACTCGGTCACTCTGAGAACCA 60
QY 61 GCCTGAGCGCTCTGTCTCACTCGGGGTAGACCAAGCGGGAGCGAGCAAGATGCGCAGA 120
DB 61 GCCTGAGCGCTCTGTCTCACTCGGGGTAGACCAAGCGGGAGCGAGCAAGATGCGCAGA 120
QY 121 CGCAGGGCAACCCGGAGGAAAGTCTGTACTACTACGAGGGGAGTGTGAAATTAATTA 180
DB 121 CGCAGGGCAACCCGGAGGAAAGTCTGTACTACTACGAGGGGAGTGTGAAATTAATTA 180
QY 181 ATGAGCAAGGCGACCCCAATGAAAGCTCAACGAAATCCGATGACTCATATTTGCTGTCA 240
DB 181 ATGAGCAAGGCGACCCCAATGAAAGCTCAACGAAATCCGATGACTCATATTTGCTGTCA 240
QY 241 ACTATGCTCTCTACCGAAAAATGAAATCTATGSCCTTCACAAAAGCAATGCTGAGAGA 300
DB 241 ACTATGCTCTCTACCGAAAAATGAAATCTATGSCCTTCACAAAAGCAATGCTGAGAGA 300
QY 301 TGACCAAGTACCAAGGAGTACATTAATTTCTGGCTGATCCCTCGAGATACA 360
DB 301 TGACCAAGTACCAAGGAGTACATTAATTTCTGGCTGATCCCTCGAGATACA 360
QY 361 TGTGCGAGTACAGAGAGATGAGATTAACAAGTGTGAGAGATCTGTCAGTATTCG 420
DB 361 TGTGCGAGTACAGAGAGATGAGATTAACAAGTGTGAGAGATCTGTCAGTATTCG 420
QY 421 ATGCGCTGTTGAGTCTGTCAAGTGTCTACTGTGTGCTGTGCGAAATGCTGTGAAC 480
DB 421 ATGCGCTGTTGAGTCTGTCAAGTGTCTACTGTGTGCTGTGCGAAATGCTGTGAAC 480
QY 481 TTATATAGCAGCAACGAGATTCGCCGTGAATTTGGGCTGGGGGCTGCACATTCAGAGA 540
DB 481 TTATATAGCAGCAACGAGATTCGCCGTGAATTTGGGCTGGGGGCTGCACATTCAGAGA 540
QY 541 AGTCGAGGCACTTGGCTTCTGTACGTCAATGATATGCTTGGCCATCTGGAACCTGC 600
DB 541 AGTCGAGGCACTTGGCTTCTGTACGTCAATGATATGCTTGGCCATCTGGAACCTGC 600
QY 601 TAAAGTATCACAGAGGCTGTGATATGATGATTAATTCACCATGTGTGACGCGCTGG 660
DB 601 TAAAGTATCACAGAGGCTGTGATATGATGATTAATTCACCATGTGTGACGCGCTGG 660
QY 661 AAGAGGCTTTTAAACCAACGAGACCGGGGTCAATGATGTCTTTCAATAGATAGAGAGT 720
DB 661 AAGAGGCTTTTAAACCAACGAGACCGGGGTCAATGATGTCTTTCAATAGATAGAGAGT 720
QY 721 ACTTCCAGAACTGGGAGACTTACGAGATACCGGGGCTGGCAAGACAGATTAATGCTG 780
DB 721 ACTTCCAGAACTGGGAGACTTACGAGATACCGGGGCTGGCAAGACAGATTAATGCTG 780
QY 781 TTAATACTCCGCTCCGAGACGGAGTTGATGAGATCTTATGAGGCCATTTTCAAGCCGG 840
DB 781 TTAATACTCCGCTCCGAGACGGAGTTGATGAGATCTTATGAGGCCATTTTCAAGCCGG 840

```

```

DB 781 TTAATACTCCGCTCCGAGACGGAGTTGATGAGATCTTATGAGGCCATTTTCAAGCCGG 840
QY 841 TCATGTCGAAGATATGAGATGATGTCAGGCTGAGGGTGTCTTACAGATGTGCTCAG 900
DB 841 TCATGTCGAAGATATGAGATGATGTCAGGCTGAGGGTGTCTTACAGATGTGCTCAG 900
QY 901 ACTCCCTATCTGGGAGTCGGTTAGGTTGCTTCAATCTATCTATCAAGACACGCCAAGT 960
DB 901 ACTCCCTATCTGGGAGTCGGTTAGGTTGCTTCAATCTATCTATCAAGACACGCCAAGT 960
QY 961 GTGTGGAATTTTGCAAGAGCTTTAACTGCTATGCTGATAGCTGGAGAGGGGTGTTACA 1020
DB 961 GTGTGGAATTTTGCAAGAGCTTTAACTGCTATGCTGATAGCTGGAGAGGGGTGTTACA 1020
QY 1021 CCATGCTGAACGTTGCCCGGTGTGACATATGAGACAGCTGTGGCCCTGATACGAGA 1080
DB 1021 CCATGCTGAACGTTGCCCGGTGTGACATATGAGACAGCTGTGGCCCTGATACGAGA 1080
QY 1081 TCCCTAATGAGCTTCCATATGACTACTTTGAATPACTTTGAGACCAAGATTTCAAGCTCC 1140
DB 1081 TCCCTAATGAGCTTCCATATGACTACTTTGAATPACTTTGAGACCAAGATTTCAAGCTCC 1140
QY 1141 ACATCAAGCTCTTCCATATGACTATACCAAGACAGATGATACCTGAGAAATCAAAC 1200
DB 1141 ACATCAAGCTCTTCCATATGACTATACCAAGACAGATGATACCTGAGAAATCAAAC 1200
QY 1201 AGGACCTGTTGAGAACCTTGAAGATGCTGCGCAGCAGCACTGTGGGCTCCAAACGAGCGA 1260
DB 1201 AGGACCTGTTGAGAACCTTGAAGATGCTGCGCAGCAGCACTGTGGGCTCCAAATGACGCGA 1260
QY 1261 TTCCTGAGAGCGCCATCTCCCTGAGAGAGTGGCGATAGAGACGAAGACCTTGACAAAGC 1320
DB 1261 TTCCTGAGAGCGCCATCTCCCTGAGAGAGTGGCGATAGAGACGAAGACCTTGACAAAGC 1320
QY 1321 GCATCTCGATCTGCTCTCTGACAAACGAATTCCTGTGAGAAAGATTTCCGATTTCTG 1380
DB 1321 GCATCTCGATCTGCTCTCTGACAAACGAATTCCTGTGAGAAAGATTTCCGATTTCTG 1380
QY 1381 AAGAGGAGGAGAGGGGGGCGCAAGAACCTTCCAACTTCAAAAAAGCCAAAGAGTCA 1440
DB 1381 AAGAGGAGGAGAGGGGGGCGCAAGAACCTTCCAACTTCAAAAAAGCCAAAGAGTCA 1440
QY 1441 AAACAGAGATGAAAAAGAAAGAACCCAGAGAGAGAAAGAAATCACCGAAGAGAGA 1500
DB 1441 AAACAGAGATGAAAAAGAAAGAACCCAGAGAGAGAAAGAAAGTCAACCGAAGAGAGA 1500
QY 1501 AAACCAAGAGAGAGAACCCAGAAAGCCAAAGGGGTCAAGAGAGGCCCAAGTGGCCCTGAA 1560
DB 1501 AAACCAAGAGAGAGAACCCAGAAAGCCAAAGGGGTCAAGAGAGGCCCAAGTGGCCCTGAA 1560
QY 1561 TGAGACCTTCACGCTCTGCTGCTGCTGAGTCCCTCAGCTTCTTTCCCTC 1611
DB 1561 TGAGACCTTCACGCTCTGCTGCTGCTGAGTCCCTCAGCTTCTTTCCCTC 1611

RESULT 4
BC000301      2093 bp      mRNA      linear      PRI 31-JAN-2003
LOCUS      Homo sapiens, similar to histone deacetylase 1, clone MGC:8378
DEFINITION      IMAGE:2820260, mRNA, complete cds.
ACCESSION      BC000301
VERSION      BC000301.1 GI:12653070
KEYWORDS      MGC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2093)
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

REMARK
COMMENT

USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaphs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akher, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Bialesley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastric, S.D., McElroy, J.C.,
McDowell, J., Pearson, R., Stantipop, S., Thomas, P.D., Touchman, J.W.,
Tsugeon, C., Vogt, J.D., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRL Plate: 1 Row: c Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13128859.

FEATURES

source

1.2093
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:8378 IMAGE:2820260"
/issue_type="lung, small cell carcinoma"
/clone_id="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
44..1492
/product="similar to histone deacetylase 1"
/protein_id="AAH0301.1"
/db_xref="GI:12653071"
/translocation="MAOTGTRKVCYVYDGVNYYYQGHMKPHRIRMTNLLN
YGLYRMEIYRPHKNAEMETKXSDYIKPLRSIRPDMSSEKMOFENGECPV
PGLPEFCOLSTGSGVSAVANKOOTDIAVMAAGIHHAKXSPSGCYVNDIYAI
LELAKHQRVLYIDIDIHGDVBEAFPTTDRVMTYSFKRGYRPGTDLIDIGAK
GRYVAVNYPLRIGIDIDESYEALEFKVMSVMEFQPSAVLDCGSDLSIDBLGCPNL
T1KGNAKCVFVKSFMIPMLMGGGYTRNVARCMYVETVALDTEIDNEPLNDYF
EYFGDPFKLHISPSMNTONTNTEYLEKIKQRLFENLRMLPHAGVOMQAIPEDAIEE
SGDEDEDDPKRISICSDKRIACDEEFSDSEEBGGRKNSNFKKARVTEDEKE
KDPBEKTEYERKEKPEAKGYEVEKLA"

CDS

BASE COUNT 570 a 480 c 531 g 512 t
ORIGIN

Query Match 93.1%; Score 1499.6; DB 9; Length 2093;
Best local similarity 99.4%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 97 GGGAGGCGAGCAAGATGGCGCAGACGCGAGGCAAGTCTGTACTACTACG 156
DB 30 GGGAGGCGAGCAAGATGGCGCAGACGCGAGGCAAGTCTGTACTACTACG 89
QY 157 ACGGGATGTTGAATTTACTATTATGAGCAAGGCCAATGAAGCCTCAACCGATCC 216
DB 90 ACGGGATGTTGAATTTACTATTATGAGCAAGGCCAATGAAGCCTCAACCGATCC 149
QY 217 GCATGACTATATTTGCTGCTCACTATGCTCTTACCGAAAAATGAAATCTATCGCC 276
DB 150 GCATGACTATATTTGCTGCTCACTATGCTCTTACCGAAAAATGAAATCTATCGCC 209
QY 277 CTCACAAAGCAATGCTGAGAGATGACCAAGTACCAAGCGATGACTACATTAAATCTT 336
DB 210 CTCACAAAGCAATGCTGAGAGATGACCAAGTACCAAGCGATGACTACATTAAATCTT 269

QY 337 TCGCTCCATCCGTCAGATTAATCATGTCGAGTACAGAGAGATGACAGATTTCAACG 396
DB 270 TCGCTCCATCCGTCAGATTAATCATGTCGAGTACAGAGAGATGACAGATTTCAACG 329
QY 397 TTGGTGAAGACTGTCAGATTTGATGAGCCTGTTTGAAGTCTGTCAATTTGTTACTGCTG 456
DB 330 TTGGTGAAGACTGTCAGATTTGATGAGCCTGTTTGAAGTCTGTCAATTTGTTACTGCTG 389
QY 457 GTTCTGTGCAAGTGTGTAAGCTTAATTAAGCAGACAGAGATCCGCGTAATGGG 516
DB 390 GTTCTGTGCAAGTGTGTAAGCTTAATTAAGCAGACAGAGATCCGCGTAATGGG 449
QY 517 CTGGGGGCTGACCATGCAAGAAAGTCCGAGGCATCTGCTTCTGTTACGTCATATGATA 576
DB 450 CTGGGGGCTGACCATGCAAGAAAGTCCGAGGCATCTGCTTCTGTTACGTCATATGATA 509
QY 577 TCGTCTTGCCCATCTGGAACCTGTAAGATCACACAGAGGCTGTACATTGACATTG 636
DB 510 TCGTCTTGCCCATCTGGAACCTGTAAGATCACACAGAGGCTGTACATTGACATTG 569
QY 637 ATATTCACATGATGACGCGCTGGAAGAGCGCTTCTACACAGAGCGGATGATGACTG 696
DB 570 ATATTCACATGATGACGCGCTGGAAGAGCGCTTCTACACAGAGCGGATGATGACTG 629
QY 697 TGTCTTTTCAATAGTATGAGAGATCTTCCAGAACTGGGGACCTTACGGGATACCGGGG 756
DB 630 TGTCTTTTCAATAGTATGAGAGATCTTCCAGAACTGGGGACCTTACGGGATACCGGGG 689
QY 757 CTGGCAAAAGCAAGTATTAATGCTGTTAACTACCCGCTCCGAGACGGATGATGACGAGT 816
DB 690 CTGGCAAAAGCAAGTATTAATGCTGTTAACTACCCGCTCCGAGACGGATGATGACGAGT 749
QY 817 CCTATGAAGGCCATTTTCAAGCCGCTCATGCTCAAGTAAAGTAAAGTTCACGCTAGT 876
DB 750 CCTATGAAGGCCATTTTCAAGCCGCTCATGCTCAAGTAAAGTAAAGTTCACGCTAGT 809
QY 877 CGGTGATCTTACAGTGTGCTCAGACTCCCTATCTGGGATGGTTAGGTTGCTTCAATC 936
DB 810 CGGTGATCTTACAGTGTGCTCAGACTCCCTATCTGGGATGGTTAGGTTGCTTCAATC 869
QY 937 TATCTATCAAGAGACACGCAAGTGTGGAATTTGTCAAGGCTTTAACTGCTGATGC 996
DB 870 TATCTATCAAGAGACACGCAAGTGTGGAATTTGTCAAGGCTTTAACTGCTGATGC 929
QY 997 TATGCTGAGAGCGGTGTTACACATTTCTTAAAGTTCGCGGTCTGAGATATGAGA 1056
DB 930 TATGCTGAGAGCGGTGTTACACATTTCTTAAAGTTCGCGGTCTGAGATATGAGA 989
QY 1057 GAGCTGTGGCCTGTGATACGAGATCCCTAATGAGCTTCCATACATGACTCTTGAAT 1116
DB 990 GAGCTGTGGCCTGTGATACGAGATCCCTAATGAGCTTCCATACATGACTCTTGAAT 1049
QY 1117 ACTTTGACCAAGATTTCAAGCTCCACATCAGTCTTCCAAATATGACTAAACAGAACGAG 1176
DB 1050 ACTTTGACCAAGATTTCAAGCTCCACATCAGTCTTCCAAATATGACTAAACAGAACGAG 1109
QY 1177 ATGAGTACTGGAAGAAAGATCAAAACAGCACTGTTTGAAGAACTTATGATGCTCCGACAG 1236
DB 1110 ATGAGTACTGGAAGAAAGATCAAAACAGCACTGTTTGAAGAACTTATGATGCTCCGACAG 1169
QY 1237 CACTGTGGGTCGAAAACGAGCGGATCTCTGAGAGACGCACTCCCTGAGAGAGTGGCGATG 1296
DB 1170 CACTGTGGGTCGAAAACGAGCGGATCTCTGAGAGACGCACTCCCTGAGAGAGTGGCGATG 1229
QY 1297 AGGAGAAAGACACCTCTGACAAAGCGATCTGATCTGCTTGAACAAAGATTTGCT 1356
DB 1230 AGGAGAAAGACACCTCTGACAAAGCGATCTGATCTGCTTGAACAAAGATTTGCT 1289
QY 1357 GTGAGAAAGATTTCTCCGATTTCTGAAGAGAGAGAGAGGGGGCGGCAAGAACTCTTCCA 1416
DB 1290 GTGAGAAAGATTTCTCCGATTTCTGAAGAGAGAGAGAGGGGGCGGCAAGAACTCTTCCA 1349

QY	1417	ACTTCAAAAAAGCCCAAGAGAGTCAAAA	CAGAGATGAAAAAGAAAAAGACCCAGAGAGA	1476
Db	1350	ACTTCAAAAAAGCCCAAGAGAGTCAAAA	CAGAGATGAAAAAGAAAAAGACCCAGAGAGA	1409
QY	1477	AGAAAGGAAATCACCAGAAAGAGAGAAA	CCAAAGAGAGAGAACCCAAAGCCAAAGGGGTCA	1536
Db	1410	AGAAAGAAAGTCACCGAAAGAGAGAAAA	CCAAAGAGAGAGAAACCCAAAGCCAAAGGGGTCA	1469
QY	1537	AGAGAGAGGCCAAGTTGGCTTGAAATGAC	CTTCTCCAGCTCTGGGCTTCTGCTGAGTCCCT	1596
Db	1470	AGGAGAGAGGTCAAGTTGGCTTGAAATGAC	CTCTCCAGCTCTGGGCTTCTGCTGAGTCCCT	1529
QY	1597	CACGTTCTTCCC	1610	
Db	1530	CACGTTCTTCCC	1543	
RESULT 5				
D50405				
LOCUS	D50405	2091 bp	mRNA	linear
DEFINITION				PRI 10-FEB-1999
ACCESSION	D50405			
VERSION	D50405.1	GI:1665722		
KEYWORDS	RPD3 protein.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Furukawa, Y., Kawakami, T., Sudo, K., Inazawa, J., Matsumine, A., Akiyama, T. and Nakamura, Y.			
TITLE	Isolation and mapping of a human gene (RPD3L1) that is homologous to RPD3, a transcription factor in <i>Saccharomyces cerevisiae</i>			
JOURNAL	Cyogenet. Cell Genet. 73 (1-2), 130-133 (1996)			
MEDLINE	96244606			
PUBMED	8646880			
REFERENCE	2 (bases 1 to 2091)			
AUTHORS	Nakamura, Y.			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 2091)			
AUTHORS	Nakamura, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-APR-1995) Yusuke Nakamura, Institute of Medical Science, The University of Tokyo, Laboratory of Molecular Medicine, Human Genome Center; 4-6-1 Shirokanebashi, Minato-ku, Tokyo 108-8639, Japan (E-mail: y-daigo@ims.u-tokyo.ac.jp, Tel:03-5449-5372, Fax:03-5449-5433)			
FEATURES				
Source	Location/Qualifiers			
	1..2091			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/chromosome="1"			
	/map="1p34.1"			
	/clone="hRPD3"			
	/tissue_type="fetal lung"			
	64..1512			
	/codon_start=1			
	/product="RPD3 protein"			
	/protein_id="BAA08909.1"			
	/db_xref="GI:1665723"			
	/translation="MAQTQGRRCVCCYYDGVNYYGQGHPPKPHRIRMTNLLINLNYGLRKMEIYRPHKANAEMTKHSDYIKELRSIRPDNMEYKQMRFNVGDCBPFVFDGFEFCOLSTGGSVAAYKLKQOQDIAVNMAGLHAKKSASGCVYNDIVLALFEDPFLGATISPNMTNQNNTNNEVLEKIKRLFENLRLPAPAGVQQAIPDAIPBSGDEDDPDKRISICSSDKRIACEEESDSEEGEGRKNSNFKAKVXTDEKEREKQDEKKEVTEEEKTKKEPKAKVKEVKLA"			
CDS				
	polya_signal			
	552 a	485 c	544 g	510 t
BASE COUNT				
ORIGIN				

	Query Match	93.0%	Score 1498;	DB 9;	Length 2091;	
	Best Local Similarity	99.3%;	Pred. No. 0;			
	Match 1504;	Conservative	0;	Mismatches	10;	Indels
						Gaps 0;
QY	97	GGAGGCGCAGCAAGATGGCCGACAGCGAGGCGACCCGGAGAAAGTCTGTACTACTACG	156			
DB	50	GGAGGCGCAGCAAGATGGCCGACAGCGAGGCGACCCGGAGAAAGTCTGTACTACTACG	109			
QY	157	ACGGGAGATGTGGAAATTACTATTATGAGCAAGGCCACCCAAATGAAGCTCACCGAATCC	216			
DB	110	ACGGGAGATGTGGAAATTACTATTATGAGCAAGGCCACCCAAATGAAGCTCACCGAATCC	169			
QY	217	GCATGACTCATAAATTTGGTCTCAACTATGCTCTCTACCGAAAAATGAAATCTATCGCC	276			
DB	170	GCATGACTCATAAATTTGGTCTCAACTATGCTCTCTACCGAAAAATGAAATCTATCGCC	229			
QY	277	CTCACAAAAGCCAAATGCTGAGAGATGACCAAGATCCACAGCATGATCTACTTAATTTCT	336			
DB	230	CTCACAAAAGCCAAATGCTGAGAGATGACCAAGATCCACAGCATGATCTACTTAATTTCT	289			
QY	337	TGGCGTCCATCCGTCCAGATATACATGTCCGAGATCAGCAAGCAGATGCGAGATTCAACG	396			
DB	290	TGGCGTCCATCCGTCCAGATATACATGTCCGAGATCAGCAAGCAGATGCGAGATTCAACG	349			
QY	397	TTGGTGAAGACTGTCCAGATTTGATGATGGCTGTTGAGTTCTGTCAATTGTCTA	456			
DB	350	TTGGTGAAGACTGTCCAGATTTGATGATGGCTGTTGAGTTCTGTCAATTGTCTA	409			
QY	457	GTTCTGAGGCAAGTGTCTGTGAATCTTAATTAAGCAGCAGAGGAGATGGCGGTAAATGGG	516			
DB	410	GTTCTGAGGCAAGTGTCTGTGAATCTTAATTAAGCAGCAGAGGAGATGGCGGTAAATGGG	469			
QY	517	CTGGGAGGCTGCAACCATGCAAGAAGTCCAGGAGATCTGGCTTCTGTATCGTCAATGATA	576			
DB	470	CTGGGAGGCTGCAACCATGCAAGAAGTCCAGGAGATCTGGCTTCTGTATCGTCAATGATA	529			
QY	577	TGCTCTTGGCCATCCTTGGAATCTGCTAAAGTATCAACCAAGGCTGTACTTGAATTTG	636			
DB	530	TGCTCTTGGCCATCCTTGGAATCTGCTAAAGTATCAACCAAGGCTGTACTTGAATTTG	589			
QY	637	ATATTACCAATGTGTGACGGCGGTGGAAGGGCTTCTACACCAACGGAACCGGGTCAATGACTG	696			
DB	590	ATATTACCAATGTGTGACGGCGGTGGAAGGGCTTCTACACCAACGGAACCGGGTCAATGACTG	649			
QY	697	TGTCCTTTCAATAGTATGAGAGTACTTCCAGGAACCTGGGAGCTTACCGGATACCGGGG	756			
DB	650	TGTCCTTTCAATAGTATGAGAGTACTTCCAGGAACCTGGGAGCTTACCGGATACCGGGG	709			
QY	757	CTGGCAAAAGCAAGTATTAATGCTGTTAACTACCCGCTCCGAGACGGGATTTGATACGAGT	816			
DB	710	CTGGCAAAAGCAAGTATTAATGCTGTTAACTACCCGCTCCGAGACGGGATTTGATACGAGT	769			
QY	817	CCATATGAGGCGCATTTTCAAGCCGGTCAATGTCCAAAGTATGAGATGTTCCAGGCTATGTG	876			
DB	770	CCATATGAGGCGCATTTTCAAGCCGGTCAATGTCCAAAGTATGAGATGTTCCAGGCTATGTG	829			
QY	877	CGGTGTGCTTACATGTGTGCTCAACATCTCCATCTTGGGATTCGGTATAGTTGCTTCAATC	936			
DB	830	CGGTGTGCTTACATGTGTGCTCAACATCTCCATCTTGGGATTCGGTATAGTTGCTTCAATC	889			
QY	937	TATCATCAAAAGACACGCCCAAGTGTGGAATTTGTCAAGAGCTTTAACTCGGCTATGTC	996			
DB	890	TATCATCAAAAGACACGCCCAAGTGTGGAATTTGTCAAGAGCTTTAACTCGGCTATGTC	949			
QY	997	TGATGTGAGGAGCGGTGTGTTACACCAATGTGAACGTTGCCGGTGTCTGACATATGAGA	1056			
DB	950	TGATGTGAGGAGCGGTGTGTTACACCAATGTGAACGTTGCCGGTGTCTGACATATGAGA	1009			
QY	1057	CAGCTGTGAGCCCTTGATATACGAGATCCCTTAATGAGCTTCCATACAAATGACTACTTTGAT	1116			
DB	1010	CAGCTGTGAGCCCTTGATATACGAGATCCCTTAATGAGCTTCCATACAAATGACTACTTTGAT	1069			

QY 1117 ACTTTGACCAATTTTCAAGCTTCCACATCACTTCTTCCAAATATGACTTAACCAAGACACA 1176
DB 1070 ACTTTGACCAATTTTCAAGCTTCCACATCACTTCTTCCAAATATGACTTAACCAAGACACA 1129
QY 1177 ATGAGTACTGGAAGATCAACAGCGACTGTTTGAACCTTAGAAGCTTAGCGCCGACG 1236
DB 1130 ATGAGTACTGGAAGATCAACAGCGACTGTTTGAACCTTAGAAGCTTAGCGCCGACG 1189
QY 1237 CACTGGGGTCCAAACGACGAGCGATTCCTGAGACGCCATCCCTGAGAGAGTGGCGATG 1296
DB 1190 CACTGGGGTCCAAACGACGAGCGATTCCTGAGACGCCATCCCTGAGAGAGTGGCGATG 1249
QY 1297 AGACGAGACGACCTCTGACAAAGCGATTCGATCTGCTCTCTGACAAACGATTCCT 1356
DB 1250 AGACGAGACGACCTCTGACAAAGCGATTCGATCTGCTCTCTGACAAACGATTCCT 1309
QY 1357 GTGAGGAAGATTTCTCCGATTTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1416
DB 1310 GTGAGGAAGATTTCTCCGATTTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
QY 1417 ACTTCAAAAAGCCAGAGAGAGTCAAAAACAGAGATGAAAAAGAAAAGAACCCAGAGAGA 1476
DB 1370 ACTTCAAAAAGCCAGAGAGAGTCAAAAACAGAGATGAAAAAGAAAAGAACCCAGAGAGA 1429
QY 1477 AGAAAGAAATCAACGAAAGAGAGAAAAACCAAGAGAGAGAACCCAGAGAACGAGGTCA 1536
DB 1430 AGAAAGAAATCAACGAAAGAGAGAAAAACCAAGAGAGAGAACCCAGAGAACGAGGTCA 1489
QY 1537 AGAGAGAGCCAAAGTTGGCTGATGATGACTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1596
DB 1490 AGAGAGAGGTCAAGTTGGCTGATGATGACTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1549
QY 1597 CAGGTTCTTTCC 1610
DB 1550 CAGGTTCTTTCC 1563

RESULT 6
AR012007
LOCUS AR012007 2111 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5763182.
ACCESSION AR012007
VERSION AR012007.1 GI:3969997
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2111)
AUTHORS Nakamura, Y. and Furukawa, Y.
TITLE RPD, protein and DNA encoding the same
JOURNAL Patent: US 5763182-A 2 09-JUN-1998;
FEATURES
source 1. 2111
location/Qualifiers
BASE COUNT 572 a 485 c 544 g 510 t
ORIGIN

Query Match 93.0%; Score 1498; DB 6; Length 2111;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 97 GGGAGGCGAGCAAGATGCGCAGACCGCAGGCAACCGGAGGAAAGTCTGTTACTACTAG 156
DB 50 GGGAGGCGAGCAAGATGCGCAGACCGCAGGCAACCGGAGGAAAGTCTGTTACTACTAG 109
QY 157 ACCGGAGTGTGAAATTAATCTATTATGACAAAGCCACCCCAATGAAGCTTCAACGATCC 216
DB 110 ACCGGAGTGTGAAATTAATCTATTATGACAAAGCCACCCCAATGAAGCTTCAACGATCC 169
QY 217 GCATGACTATATTTGCTGCTCACTATGCTCTTACCGAAAAATGGAATCTATCGCC 276
DB 170 GCATGACTATATTTGCTGCTCACTATGCTCTTACCGAAAAATGGAATCTATCGCC 229

QY 277 CTCACAAAGCCATATGCTGAGAGATGACCAAGTACCACAGCATGACTATTAATTTCT 336
DB 230 CTCACAAAGCCATATGCTGAGAGATGACCAAGTACCACAGCATGACTATTAATTTCT 289
QY 337 TGGGCTCCATCCGTCAGATTAATGTCGGAATTAACAGAGAGATGAGATTCACG 396
DB 290 TGGGCTCCATCCGTCAGATTAATGTCGGAATTAACAGAGAGATGAGATTCACG 349
QY 397 TTGTGAGAGACTGTCAGATTCGATGAGCTGTTGAGTTCTGTGAGTTGTCAGTTGTCAGTGTG 456
DB 350 TTGTGAGAGACTGTCAGATTCGATGAGCTGTTGAGTTCTGTGAGTTGTCAGTTGTCAGTGTG 409
QY 457 GTTCTGTGGAAGTCTGTGAATTTAATAGCAGACGACGACATTCGCTGAATTTGG 516
DB 410 GTTCTGTGGAAGTCTGTGAATTTAATAGCAGACGACGACATTCGCTGAATTTGG 469
QY 517 CTGGGGGCTTGACCATGCAAGAAAGTCCGAGGCAATTCGCTTCTGTACGTCATAGATA 576
DB 470 CTGGGGGCTTGACCATGCAAGAAAGTCCGAGGCAATTCGCTTCTGTACGTCATAGATA 529
QY 577 TCGTCTGGCCATCTCGAACTGTAAAGTATCACAGAGGCTGTACATTGACATTG 636
DB 530 TCGTCTGGCCATCTCGAACTGTAAAGTATCACAGAGGCTGTACATTGACATTG 589
QY 637 ATATTCACCATGATGACGCGCTGAAAGAGCCCTTCTACACAGACCGGATCATGACTG 696
DB 590 ATATTCACCATGATGACGCGCTGAAAGAGCCCTTCTACACAGACCGGATCATGACTG 649
QY 697 TGTCTTTTATATATATGAGAGATTTCTCCAGAACTGGGGACCTTACCGGATATCCGGGG 756
DB 650 TGTCTTTTATATATATGAGAGATTTCTCCAGAACTGGGGACCTTACCGGATATCCGGGG 709
QY 757 CTGGCAAAAGCAAGTATTTATGCTGTTAACTAACCGCTCCGAGACGGGATTTGATGACGAGT 816
DB 710 CTGGCAAAAGCAAGTATTTATGCTGTTAACTAACCGCTCCGAGACGGGATTTGATGACGAGT 769
QY 817 CCTATGAGGCCATTTTCAAGCCGATCATGTCCAAAGTAAATGAGATGTTCCAGCCTAGTG 876
DB 770 CCTATGAGGCCATTTTCAAGCCGATCATGTCCAAAGTAAATGAGATGTTCCAGCCTAGTG 829
QY 877 CGGTGCTTTAAGATGTCGCTCAAGTCTCTTATCTGGGGATGCTTAAAGTCTTAAATC 936
DB 830 CGGTGCTTTAAGATGTCGCTCAAGTCTCTTATCTGGGGATGCTTAAAGTCTTAAATC 889
QY 937 TATCTATCAAAAGACACGCGCAAGTGTGGAATTTGTCAAGAGCTTTAATCTGCTATGC 996
DB 890 TATCTATCAAAAGACACGCGCAAGTGTGGAATTTGTCAAGAGCTTTAATCTGCTATGC 949
QY 997 TGATGCTGGAGCGGTGTTACACCATTCGTACGTTGCCGTCGTCGATATGAGA 1056
DB 950 TGATGCTGGAGCGGTGTTACACCATTCGTACGTTGCCGTCGTCGATATGAGA 1009
QY 1057 CAGCTGTGCTTGAATACGAGATTCCTAATGAGCTTCAATACATGACTACTTTGAAT 1116
DB 1010 CAGCTGTGCTTGAATACGAGATTCCTAATGAGCTTCAATACATGACTACTTTGAAT 1069
QY 1117 ACTTTGACCAATTTTCAAGCTTCCACATCACTTCTTCCAAATATGACTTAACCAAGACACA 1176
DB 1070 ACTTTGACCAATTTTCAAGCTTCCACATCACTTCTTCCAAATATGACTTAACCAAGACACA 1129
QY 1177 ATGAGTACTGGAAGATCAACAGCGACTGTTTGAACCTTAGAAGCTTAGCGCCGACG 1236
DB 1130 ATGAGTACTGGAAGATCAACAGCGACTGTTTGAACCTTAGAAGCTTAGCGCCGACG 1189
QY 1237 CACTGGGGTCCAAACGACGAGCGATTCCTGAGACGCCATCCCTGAGAGAGTGGCGATG 1296
DB 1190 CACTGGGGTCCAAACGACGAGCGATTCCTGAGACGCCATCCCTGAGAGAGTGGCGATG 1249
QY 1297 AGACGAGACGACCTCTGACAAAGCGATTCGATCTGCTCTCTGACAAACGATTCCT 1356
DB 1250 AGACGAGACGACCTCTGACAAAGCGATTCGATCTGCTCTCTGACAAACGATTCCT 1309
QY 1357 GTGAGGAAGATTTCTCCGATTTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1416

Db 1310 GTGAGGAGAGTTCTCCGATCTCTAGAGGAGGAGGAGGCGGCGCAGAAAGCTTCCCA 1369
Qy 1417 ACTTCAAAAAAGCCAGAGAGTCAAAA CAGAGATGAAAAAGAAAAGACCAGAGAGA 1476
Db 1370 ACTTCAAAAAAGCCAGAGAGTCAAAA CAGAGATGAAAAAGAAAAGACCAGAGAGA 1429
Qy 1477 AGAAAGATCAAGCCAGAGAGAGAAAA CCAAGAGAGAAAGCCAGAAAGGGGTCA 1536
Db 1430 AGAAAGAGTCAAGCCAGAGAGAGAAAA CCAAGAGAGAAAGCCAGAAAGGGGTCA 1489
Qy 1537 AGAGAGAGCCAGAGTGGCTGGAATGGA CTTCTCAGCTCTGCTTCTGCTGATCCCT 1596
Db 1490 AGAGAGAGTCAAGTGGCTGGAATGGA CTTCTCAGCTCTGCTTCTGCTGATCCCT 1549
Qy 1597 CAGCTTCTTCCC 1610
Db 1550 CAGCTTCTTCCC 1563

RESULT 7
E11455
LOCUS E11455 2111 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human transcriptional control protein, RPDL.
ACCESSION E11455
VERSION E11455.1 GI:22025091
KEYWORDS JP 1996140687-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2111)
AUTHORS Nakamura, Y. and Furukawa, Y.
TITLE RPDL PROTEIN AND DNA ENCODING FOR THE SAME
JOURNAL PATENT: JP 1996140687-A 1 04-JUN-1996;
JAPAN FOUND CANCER RES; EISAI CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996140687-A/1
PD 04-JUN-1996
PF 20-JUL-1995 JP 1995183763
PI 22-SEP-1994 JP 94P 227876
PT NAKAMURA YOSUKE, FURUKAWA YOICHI
PC C121N15/09, C07K14/39, C07K16/18, C12P21/02, C12P21/08, C12Q1/68, PC
(C12P21/02,
C12R1:19);
PC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2111
FT /organism="Homo sapiens"
FT /tissue="Lung"
FT /clone="L1-3977"
FT 5'UTR 1..63
FT CDS 64..1512
FT /product="human transcriptional control protein, RPDL"
FT 3'UTR 1513..2111.
FEATURES
source Location/Qualifiers
1..2111
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
BASE COUNT 572 a 485 c 544 g 510 t
ORIGIN
Query Match 93.0%; Score 1498; DB 6; Length 2111;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 97 GGGAGGCGAGCAAGATGGCGCAGACGACAGGCGACCCGAGAGAAAGTCTGTTACTACTACG 156

Db 50 GGGAGGCGAGCAAGATGGCGCAGACGACGCGACCCGAGAGAAAGTCTGTTACTACTACG 109
Qy 157 ACCGGGAGTGTGAAATTTACTATTATGACAAAGGCCACCAATGAAGCCCTCACGAAATCC 216
Db 110 ACCGGGAGTGTGAAATTTACTATTATGACAAAGGCCACCAATGAAGCCCTCACGAAATCC 169
Qy 217 GCATGACTCAATATTTGCTGCTCACTAGTGTCTTACCCGAAAAATGAAATATATGCCC 276
Db 170 GCATGACTCAATATTTGCTGCTCACTAGTGTCTTACCCGAAAAATGAAATATATGCCC 229
Qy 277 CTGACAAAGCCATGCTGAGAGATGACCAAGTACCAAGCATGACTACTAATTAATTTCT 336
Db 230 CTGACAAAGCCATGCTGAGAGATGACCAAGTACCAAGCATGACTACTAATTAATTTCT 289
Qy 337 TGGGCTCCATCGTCCAGATTAATGCGAGTACAGCAAGAGAGAGAGATTCACG 396
Db 290 TGGGCTCCATCGTCCAGATTAATGCGAGTACAGCAAGAGAGAGAGATTCACG 349
Qy 397 TTGGTGAAGACTGTCCAGATTAATGCGAGTACAGCAAGAGAGAGAGATTCACG 456
Db 350 TTGGTGAAGACTGTCCAGATTAATGCGAGTACAGCAAGAGAGAGAGATTCACG 409
Qy 457 GTTCTGTGCAAGTCTGTGAAATTTAATAGCAGACAGCAATCGCCGTGAATTTGGG 516
Db 410 GTTCTGTGCAAGTCTGTGAAATTTAATAGCAGACAGCAATCGCCGTGAATTTGGG 469
Qy 517 CTGGGGGCGCTGCACATGCAGAAAGAGCGAGGATCTGCTGTGATAGTCAATGATA 576
Db 470 CTGGGGGCGCTGCACATGCAGAAAGAGCGAGGATCTGCTGTGATAGTCAATGATA 529
Qy 577 TCGTCTTGGCCATCTGGAATCTGTAAGTATCAGAGAGGTGCTGTACATTGACATTG 636
Db 530 TCGTCTTGGCCATCTGGAATCTGTAAGTATCAGAGAGGTGCTGTACATTGACATTG 589
Qy 637 ATATTCACATGCTGACGCGGTGGAAGAGGCTTTCTACACACGAGACCGGGTCAATGACTG 696
Db 590 ATATTCACATGCTGACGCGGTGGAAGAGGCTTTCTACACACGAGACCGGGTCAATGACTG 649
Qy 697 TGTGCTTTCTAATAGTATGAGAGTATCCAGAGAACTGGGGGACTATGAGGATACCGGGG 756
Db 650 TGTGCTTTCTAATAGTATGAGAGTATCCAGAGAACTGGGGGACTATGAGGATACCGGGG 709
Qy 757 CTGGCAAAAGCAAGTATTTATGCTTTAATACCCGCTCCAGAGCGGATTTGATGACGAGT 816
Db 710 CTGGCAAAAGCAAGTATTTATGCTTTAATACCCGCTCCAGAGCGGATTTGATGACGAGT 769
Qy 817 CTTATGAGGCAATTTTCAAGCCGCTCATGTCCAAAGTATGAGATTTCAAGCTTATGTC 876
Db 770 CTTATGAGGCAATTTTCAAGCCGCTCATGTCCAAAGTATGAGATTTCAAGCTTATGTC 829
Qy 877 CGGTGCTTCAAGTGTGCTGACAGCTCCCTATCTGGGGATCGGTTAGTGTGCTCAATC 936
Db 830 CGGTGCTTCAAGTGTGCTGACAGCTCCCTATCTGGGGATCGGTTAGTGTGCTCAATC 889
Qy 937 TATCTATCAAAAGACACGCAAGTGTGTGGAATTTGTCAAGAGCTTTAAGCTGCTATGC 996
Db 890 TATCTATCAAAAGACACGCAAGTGTGTGGAATTTGTCAAGAGCTTTAAGCTGCTATGC 949
Qy 997 TGAATGCGAGGCGGTGTTTACACCAATGTAACGTTGCGGTGTGACATATGAGA 1056
Db 950 TGAATGCGAGGCGGTGTTTACACCAATGTAACGTTGCGGTGTGACATATGAGA 1009
Qy 1057 CAGCTGTGCGCTGTGATACCGAATCTCTTATAGAGCTTCAATCAATGACTACTTGAAT 1116
Db 1010 CAGCTGTGCGCTGTGATACCGAATCTCTTATAGAGCTTCAATCAATGACTACTTGAAT 1069
Qy 1117 ACTTGGACAGATTTCAAGCTCCATCAGTCTTCCATATAGACTTAACAGAACAGCA 1176
Db 1070 ACTTGGACAGATTTCAAGCTCCATCAGTCTTCCATATAGACTTAACAGAACAGCA 1129
Qy 1177 ATGAGTACTGTGAGAAAGATCAAAACAGGACTGTTTGAACCTTGAATGCTGCGCAGC 1236

Db	1130	ATGAGTACTCGAAGAAGATCAAAACGGCAGCTGTTTGAGAACCTTAGATTGTGCGCCACG	1189
Oy	1237	CACCTGGGGGTCCAAAACGTAAGCGATTCCTTGAGACGCCATCCCTGAGAGAGTGGCGATG	1296
Db	1190	CACCTGGGGGTCCAATGTAGGCGATTCCTTGAGAGCGCATCCCTGAGAGAGTGGCGATG	1249
Oy	1297	AGAGCGAAGAAGACCCCTGACAAGGCGCATCTGCATCTGCCTCGACAAAGAAATGGCT	1356
Db	1250	AGAGCGAAGAAGACCCTGACAAAGGCGCATCTGCATCTGCCTCGACAAAGAAATGGCT	1309
Oy	1357	GTCGGAAGAAGATTCCTCCGATTCCTGAAAGAGAGGAGAGAGGAGGCGCACGAATCTTCC	1416
Db	1310	GTGAGGAAGAAGATTCCTCCGATTCCTGAAAGAGAGGAGAGAGGAGGCGCACGAATCTTCC	1369
Oy	1417	ACTTCAAAAAACCBAAGAGAGTCAAAACAGAGATGATAAAAGAGAAAGACCCAGAGAGA	1476
Db	1370	ACTTCAAAAAACCBAAGAGAGTCAAAACAGAGATGATAAAAGAGAAAGACCCAGAGAGA	1429
Oy	1477	AGAAAGGAATCACCGAABAAGAGAAACCAAAGAGAGAAACCCAGAACCCAAAGGGGTCA	1536
Db	1430	AGAAAGGAATCACCGAABAAGAGAAACCAAAGAGAGAAACCCAGAACCCAAAGGGGTCA	1489
Oy	1537	AGAGAGAGGCGCAAGTTGGCTCGAATAGACCTCTCCAGCTCTGGCTTCCTGCTGAGTCCCT	1596
Db	1490	AGAGAGAGGTCAAGTTGGCTCGAATAGACCTCTCCAGCTCTGGCTTCCTGCTGAGTCCCT	1549
Oy	1597	CACGTTTCTTTCCC 1610	
Db	1550	CACGTTTCTTCCCC 1563	
RESULT 8			
LOCUS	162388	2111 bp	DNA linear PAT 07-OCT-1997
DEFINITION	Sequence 2 from patent US 5659016.		
ACCESSION	E62388		
VERSION	E62388.1 GI:2480336		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2111)		
AUTHORS	Nakamura,Y. and Furukawa,Y.		
TITLE	RPLD protein and DNA encoding the same		
JOURNAL	Patent: US 5659016-A 2 19-MUG-1997;		
FEATURES	location/Qualifiers		
source	1..2111 /organism="unknown"		
BASE COUNT	572 a 485 c 544 g 510 t		
ORIGIN			
Query Match:	93.0%; Score 1498; DB 6; Length 2111;		
Bases Local Similarity	99.3%; Pred. No. 0; Mismatches 10; Indels 0; Gaps 0;		
Matches 1504; Conservative	0;		
Oy	97	GGGAGGCGAGCAAGATGGCGCGACGACGAGGCGACCCGGAGGAAAGTCTGTTACTACTACG	156
Db	50	GGGAGGCGAGCAAGATGGCGCGACGACGAGGCGACCCGGAGGAAAGTCTGTTACTACTACG	109
Oy	157	ACGGGGATGTTGGAAATTACTATTATGACAAGGCGACCCCAATGAAAGCTTCACGGAATCC	216
Db	110	ACGGGGATGTTGGAAATTACTATTATGACAAGGCGCACCAATGAAAGCTTCACGGAATCC	169
Oy	217	GCATGACTCATTAATTGCTGCTGCATCTATGCTCTTACCAGAAAATGGAATCTATGCGC	276
Db	170	GCATGACTCATTAATTGCTGCTGCATCTATGCTCTTACCAGAAAATGGAATCTATGCGC	229
Oy	277	CTCACAAAGCCAAATCTGAGAGAGATGACCAAGTATCACAGCGATGACTTAAATTTCT	336
Db	230	CTCACAAAGCCAAATCTGAGAGAGATGACCAAGTATCACAGCGATGACTTAAATTTCT	289
Oy	337	TGCGCTCATTCGTCCAAGTAACATGTCGAGATACAGCAAGCATGACAGATTCAAACG	396

Db	290	TGGCTCCATCCGTCAGATTAACAATGTCGGAGTACACCAAGCAGATCAGAGATTCAACG	349
OY	397	TTGGTAGAGACTGTCACGATATTGCATGGCCCTGTTAGATTCTGTCAGTTGCTTACTGGTG	455
Db	350	TTGGTAGAGACTGTCACGATATTGCATGGCCCTGTTAGATTCTGTCAGTTGCTTACTGGTG	409
OY	457	GTTCTGAGGCAAGGCTGTGAAACTTAATTAAGCAGCAGACGCAATGCGCCGTAAATGGG	516
Db	410	GTTCTGAGGCAAGGCTGTGAAACTTAATTAAGCAGCAGACGCAATGCGCTGTGAATGGG	469
OY	517	CTGGGGGCGCTGCACCATGCAAAAGAGTCCGAGGATCTGGCTTCTGTATCGTCAATGATA	576
Db	470	CTGGGGGCGCTGCACCATGCAAAAGAGTCCGAGGATCTGGCTTCTGTATCGTCAATGATA	529
OY	577	TCGTCTTGCCCATCTCGAACTGCTAAAGTATCCAGAGGGTCTGTACTTGACATTG	636
Db	530	TCGTCTTGCCCATCTCGAACTGCTAAAGTATCCAGAGGGTCTGTACTTGACATTG	589
OY	637	ATATTCACATGAGGAGCGGCTGGAAGAAGCGCTTCTACACAGGACCGGGGTCAATGACTG	696
Db	590	ATATTCACATGAGGAGCGGCTGGAAGAAGCGCTTCTACACAGGACCGGGGTCAATGACTG	649
OY	697	TGTCCTTTTCATTAAGTATGAGAGACTTCCAGAGACTGGGAGCTACGCGGATACCGGG	756
Db	650	TGTCCTTTTCATTAAGTATGAGAGACTTCCAGAGACTGGGAGCTACGCGGATACCGGG	709
OY	757	CTGGCAAAAGACAAGATTAATGCTGTAACTACCCGCTCCGAGCGGATTTGATGACGAGT	816
Db	710	CTGGCAAAAGACAAGATTAATGCTGTAACTACCCGCTCCGAGCGGATTTGATGACGAGT	769
OY	817	CCTATGAGGCCATTTTCAAGCCGCTCATGTCCTCAAGTAATGAGAAATGTTCCAGGCTTAGT	876
Db	770	CCATATGAGGCCATTTTCAAGCCGCTCATGTCCTCAAGTAATGAGAAATGTTCCAGGCTTAGT	829
OY	877	CGGAGGACTTACAGTGTGGGCTCAACATCCCATCTGGGAGTCGGTTAGTGTGCTTCAATC	936
Db	830	CGGAGGACTTACAGTGTGGGCTCAACATCCCATCTGGGAGTCGGTTAGTGTGCTTCAATC	889
OY	937	TATCTATCAAAAGACAACGCCAAGTGTGTGTAATTTGTCAAGAGCTTTAACTCGGCTATGC	996
Db	890	TAACTATCAAAAGACAACGCCAAGTGTGTGTAATTTGTCAAGAGCTTTAACTCGGCTATGC	949
OY	997	TGATGCTGGAGGCGGTGTGTTACCAATTCGTAACGTTGCCGGTGTGTGACATATGAGA	1056
Db	950	TGATGCTGGAGGCGGTGTGTTACCAATTCGTAACGTTGCCGGTGTGTGACATATGAGA	1009
OY	1057	CAGCTGTGGCCCTGTGATACGAGATTCCTCTATGAGCTTCCATAAATGATCAATCTTGAT	1116
Db	1010	CAGCTGTGGCCCTGTGATACGAGATTCCTCTATGAGCTTCCATAAATGATCAATCTTGAT	1069
OY	1117	ACTTTGACACGATTTCAAGCTCCACATCAGTCCCTTCCATATGACTTAACGAGAACAGA	1176
Db	1070	ACTTTGACACGATTTCAAGCTCCACATCAGTCCCTTCCATATGACTTAACGAGAACAGA	1129
OY	1177	ATGAGTACTTGAGAAAGATCAAAACAGCGACTGTTTGAAACCTTAGAATGCTGCCGACG	1236
Db	1130	ATGAGTACTTGAGAAAGATCAAAACAGCGACTGTTTGAAACCTTAGAATGCTGCCGACG	1189
OY	1237	CACCTGGGGGTCAAAACGACGGCAATTTCTGAGAGACGCCATCCCTGAGAGAGTGGCATG	1296
Db	1190	CACCTGGGGGTCAAAATGCAAGCGAATTTCTGAGAGACGCCATCCCTGAGAGAGTGGCATG	1249
OY	1297	AGGACGAAGACGACCCCTGACAAAGGCGATCTGATCTGCTCTGACAAACGAATTTGCT	1356
Db	1250	AGGACGAAGACGACCCCTGACAAAGGCGATCTGATCTGCTCTGACAAACGAATTTGCT	1309
OY	1357	GTTGAGGAAGATTCTCCGATTCTGAAGAGAGGAGAGGGGGGCGCAAGAACTCTTCCA	1416
Db	1310	GTTGAGGAAGATTCTCCGATTCTGAAGAGAGGAGAGGGGGGCGCAAGAACTCTTCCA	1369
OY	1417	ACTTCAAAAAAGCCAAAGAGTCAAAACAGAGATGAAAAAGAAAAAGCCCAAGAGAGA	1476
Db	1370	ACTTCAAAAAAGCCAAAGAGTCAAAACAGAGATGAAAAAGAAAAAGCCCAAGAGAGA	1429

QY	1477	AGAAAGAAATACCGAAGAGAGAAAA	CAAAGAGAGAGAGAGCC	CAAGAGCC	CAAGAGGCTCA	1536
Db	1430	AGAAAGAAATACCGAAGAGAGAAAA	CAAAGAGAGAGAGAGCC	CAAGAGGCTCA	1489	
QY	1537	AGAGAGAGAGCCAGATTGGCTGTAATGGAACCTCTTCAGCTCTGCGTTCCTGTAATGCTCT	1596			
Db	1490	AGAGAGAGAGTCAATTGGCTGTAATGGAACCTCTTCAGCTCTGCGTTCCTGTAATGCTCT	1549			
QY	1597	CACGTTCTTCTCC	1610			
Db	1550	CACGTTCTTCTCCC	1563			

BASE COUNT		405 a	330 c	395 g	319 t	SDGDDDDDDPKRISICSDSDRIACBEEFSEFSEBEGGGRKNSNFKAKRVKTEDEKE KDPEKKEVTEBEKTEBEKREAGVEHYKAL"	
ORIGIN							
Query Match	89.0%; Score 1434.2;	DB 12;	Length 1449;				
Best Local Similarity	99.4%;	Pred. No. 0;					
Matches 1439;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;			
QY	111	ATGCGCGAGACGACGGGCACTCCGAGAGAAAGTCTGTTACTACTACGACGGGAGATGTTGGA	170				
Db	1	ATGGCGCAGACGACGGGCACTCCGAGAGAAAGTCTGTTACTACTACGACGGGAGATGTTGGA	60				
QY	171	AATTACTATTATGAGACAAAGGCCCAATGAAGCCTACCGGATCCGATGACATCAAT	230				
Db	61	AATTACTATTATGAGACAAAGGCCCAATGAAGCCTACCGGATCCGATGACATCAAT	120				
QY	231	TTGCTGCTCACTATGAGTCTCTACCCGAAATATGAAATCTATCGCCCTCACAAAAGCAAT	290				
Db	121	TTGCTGCTCACTATGAGTCTCTACCCGAAATATGAAATCTATCGCCCTCACAAAAGCAAT	180				
QY	291	GCTGAGAGAGATGACCAAGTACACACGAGTACTACTAATTAATTTTGCGCTCCATCCGT	350				
Db	181	GCTGAGAGAGATGACCAAGTACACACGAGTACTACTAATTAATTTTGCGCTCCATCCGT	240				
QY	351	CCAGATAACATGTTCGAGATACAGCAGCAAGATGACAGATTTCAACGTTGGTGAAGGACTGT	410				
Db	241	CCAGATAACATGTTCGAGATACAGCAGCAAGATGACAGATTTCAACGTTGGTGAAGGACTGT	300				
QY	411	CCAGTATTTCAGTGGCCGTGTTGAGTTCTGTCAGTTCCTACTGGTGGTTCTGTGGCAAGT	470				
Db	301	CCAGTATTTCAGTGGCCGTGTTGAGTTCTGTCAGTTCCTACTGGTGGTTCTGTGGCAAGT	360				
QY	471	GCTGTGAAACTTTAATAAGACAGACGGACATCGCGGTGAATTTGGGCTCGGGGCGCTGCAC	530				
Db	361	GCTGTGAAACTTTAATAAGACAGACGGACATCGCGGTGAATTTGGGCTCGGGGCGCTGCAC	420				
QY	531	CATGCAAAAGAGTCCGAGGCATCTGGCTTTCTGTTACGTTCAATGATATACGTTCTGGGCATC	590				
Db	421	CATGCAAAAGAGTCCGAGGCATCTGGCTTTCTGTTACGTTCAATGATATACGTTCTGGGCATC	480				
QY	591	CTGGACCTGCTAAAGTATACACAGAGGGTGTGATACATTTGATATATTCACCATGCT	650				
Db	481	CTGGACCTGCTAAAGTATACACAGAGGGTGTGATACATTTGATATATTCACCATGCT	540				
QY	651	GACGCGGTGAGAAAGGCTTCTACACACCGGACCGGCTCATGACTGTGTCTTTCATPAAG	710				
Db	541	GACGCGGTGAGAAAGGCTTCTACACACCGGACCGGCTCATGACTGTGTCTTTCATPAAG	600				
QY	711	TATGAGAGATCTTCCAGGAACTTGGGGAAGCTTACGGGATACCGGGGCTGGCAAAAGACAAG	770				
Db	601	TATGAGAGATCTTCCAGGAACTTGGGGAAGCTTACGGGATACCGGGGCTGGCAAAAGACAAG	660				
QY	771	TATTATGCTGTAACTACCGGCTCCGAGACGGGATATGATGACAGTCTCTATGAGGCAATT	830				
Db	661	TATTATGCTGTAACTACCGGCTCCGAGACGGGATATGATGACAGTCTCTATGAGGCAATT	720				
QY	831	TTCAAGCCGGTCACTGTCCAAAGTATGAGATGTTCCAGCCTTAGTCGGTGTCTTTACAG	890				
Db	721	TTCAAGCCGGTCACTGTCCAAAGTATGAGATGTTCCAGCCTTAGTCGGTGTCTTTACAG	780				
QY	891	TGTGGCTCAGACTCCCTATCTGGGGATTCGGTTAGTTGCTTCAATCTATCTATCAAAAGA	950				
Db	781	TGTGGCTCAGACTCCCTATCTGGGGATTCGGTTAGTTGCTTCAATCTATCTATCAAAAGA	840				
QY	951	CACGCCAAGTGTGGAATTTGTCAAGAAGTTTAAACGCTATGCTGATGCTGGAGGC	1010				
Db	841	CACGCCAAGTGTGGAATTTGTCAAGAAGTTTAAACGCTATGCTGATGCTGGAGGC	900				
QY	1011	GGTGGTTACACATTTGTAAGCTTGGCCGGTGTCTGACATATGAGACAGCTGTGGCCCTG	1070				
Db	901	GGTGGTTACACATTTGTAAGCTTGGCCGGTGTCTGACATATGAGACAGCTGTGGCCCTG	960				


```

QY 1071 GATACGAGATCCCTTAATGAGCTTCCTCAATAGACTACTTTGAATACCTTTGGACAGAT 1130
DB 961 GATACGAGATCCCTTAATGAGCTTCCTCAATAGACTACTTTGAATACCTTTGGACAGAT 1020
QY 1131 TTCAAGCTCCCATATGAGTCTCTTCCAAATATGACTTAACCAAGACCAAGATAGTACCTGGAG 1190
DB 1021 TTCAAGCTCCCATATGAGTCTCTTCCAAATATGACTTAACCAAGACCAAGATAGTACCTGGAG 1080
QY 1191 AAGATCAACAGAGAGTCTTTTGAAGACTTGAATGCTGCGGACGACCTGGGGTCCAA 1250
DB 1081 AAGATCAACAGAGAGTCTTTTGAAGACTTGAATGCTGCGGACGACCTGGGGTCCAA 1140
QY 1251 ACGCAGGCGATTCCTGAGAGACCGCATCCCTGAGAGAGTGGCGATGAGAGCAAGACGAC 1310
DB 1141 ATGACAGGCGATTCCTGAGAGACCGCATCCCTGAGAGAGTGGCGATGAGAGCAAGACGAC 1200
QY 1311 CCTGACAGGCGATTCCTGAGAGACCGCATCCCTGAGAGAGTGGCGATGAGAGCAAGACGAC 1370
DB 1201 CCTGACAGGCGATTCCTGAGAGACCGCATCCCTGAGAGAGTGGCGATGAGAGCAAGACGAC 1260
QY 1371 TCCGATTCGAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1430
DB 1261 TCCGATTCGAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1431 AAGAGAGTCAAAACAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490
DB 1321 AAGAGAGTCAAAACAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1491 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1550
DB 1381 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1551 TTGGGCT 1557
DB 1441 TTGGGCT 1447

```

```

RESULT 10
MMHIDE 1977 bp mRNA linear ROD 09-SEP-1997
LOCUS M.musculus mRNA for histone deacetylase.
DEFINITION X98207
VERSION X98207.1 GI:1771285
KEYWORDS histone deacetylase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Bartl S., Taplick J., Lagger G., Khier H., Kuchler K. and Seiser C.
TITLE Identification of mouse histone deacetylase 1 as a growth
factor-inducible gene
JOURNAL Mol. Cell. Biol. 17 (9), 5033-5043 (1997)
MEDLINE 97415582
PUBMED 9271381
REFERENCE 2 (bases 1 to 1977)
AUTHORS Seiser C.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1996) C. Seiser, University of Vienna, Institute
of Molecular Biology, Vienna Biocenter, Dr.Bohr-Gasse 9, A-1030
Vienna, AUSTRIA
FEATURES
source 1.1977 location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/feature="Swiss 3T3"
18.1466
/codon_start=1
/product="histone deacetylase"
/protein_id="CAA66870.1"
/db_xref="GI:1771285"
/db_xref="MGI:108086"

```

```

/db_xref="SWISS-PROT:009106"
/translation="MAOTGTRKRVKCYDDVDVNGYVYGGQHPMKPHIRIRMTNLLN
YGLRKMETIYPHKNAAEMTKYSDYDKFLRSIRPNMSYSKOMRPNVGECPV
FDGJFEPCOLSTGGSVASAVLNKOQOTIIVNAGSLHAKKSEASGCVYDIYAI
LELKYHORYVYIYDIDIHGSGVEAAPTTPVWMTVSFKYGEVPRGQDLPIDGAK
KRTIAYNPFLRDGIDDESEYEAIFKWSKVMFOPSLVILQSGDSISGRBLGPNL
TIKGAIAKVEFVKSFNLMPLMGGGYIRNVARWTYETVAALDEIIPNELVNDYF
EYFGDFLHISPSNMTQNTNNEYLEKIKFLFENLRMLPHAPGVOMQAIPEDAIPEE
SGDEBEDPDKRISICSSDKRIACEEESDSDEBEGGRKNSNFKKAKRVTEDEKE
KDPBEKRVTEBEKTEKPEAKVKEEVKLA"
BASE COUNT 543 a 475 c 488 g 471 t
ORIGIN
Query Match 77.0%; Score 1240.2; DB 10; Length 1977;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1349; Conservative 0; Mismatches 163; Indels 1; Gaps 1;
QY 94 GCGGGAAGGCGAGAGATGCGGCGAGAGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 153
DB 1 GCCGGGCGGCGAGAGATGCGGCGAGAGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 154 ACGAGGGAGAGTGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 213
DB 61 ACGAGGGAGAGTGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120
QY 214 TCCGATGATCAATTAATTTGCTGCTCAATTAATTTGCTGCTCAATTAATTTGCTGCTCAATTAAT 273
DB 121 TCCGATGATCAATTAATTTGCTGCTCAATTAATTTGCTGCTCAATTAATTTGCTGCTCAATTAAT 180
QY 274 GCCCTCAAAAGCCAAATGCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 333
DB 181 GCCCTCAAAAGCCAAATGCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 334 TCTTGCGCTCCATCCGCTCAAGTAAATCAATGTCGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 393
DB 241 TCTTGCGCTCCATCCGCTCAAGTAAATCAATGTCGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 394 ACGTTGGAGAGAGTGTCCAGATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453
DB 301 ACGTTGGAGAGAGTGTCCAGATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 454 GTGCTTCTGTGCAAGTGTGTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 513
DB 361 GAGGCTGTGTGCAAGTGTGTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
QY 514 GGGCTGGGGGCTGCAATGCAAGAGAGTCCGAGAGATCTGGCTTCTGTACGTCAATG 573
DB 421 GGGCTGGGGGCTGCAATGCAAGAGAGTCCGAGAGATCTGGCTTCTGTACGTCAATG 480
QY 574 AATATGCTTGGGCAATCCGAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 633
DB 481 AATATGCTTGGGCAATCCGAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
QY 634 TTGATATTAACCATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693
DB 541 TTGATATTAACCATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 694 CTGTGCTCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 753
DB 601 CTGTGCTCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
QY 754 GGGCTGGGAAAGAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 813
DB 661 GGGCTGGGAAAGAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 814 AGTCTATAGAGCCATTTTCAAGCGGATGATCCAAAGTAAATTAATTAATTAATTAATTAATTAAT 873
DB 721 AATCTATAGAGCCATTTTCAAGCGGATGATCCAAAGTAAATTAATTAATTAATTAATTAATTAAT 780
QY 874 GTGCGGTGCTTAAAGTGTGCTCAAGTCTGAGTAAATTAATTAATTAATTAATTAATTAATTAAT 933
DB 781 GTGCGGTGCTTAAAGTGTGCTCAAGTCTGAGTAAATTAATTAATTAATTAATTAATTAATTAAT 840

```

QY 934 ATCTATATCAAAAGACACGCCAAGTGTGTGAATTTGTCAAGAGCTTTAACTCTGCTA 993
 DB 841 ATCTGACCATCAAAAGACACGCCAAGTGTGTGAATTTGTGAAGTTTCAACTTTGCCA 900
 QY 994 TGTGATGCTGGAGGCGGTGTGTACACCATTTGTTAACTTTGCCGCTGTGACATATG 1053
 DB 901 TGTGATGCTGGAGGCGGTGTGTACACCATTTGTTAACTTTGCCGCTGTGACATATG 960
 QY 1054 AGACAGCTGGGCGCTGGATACGAGATCCCTATGAGCTTCCATACATGACTACTTTG 1113
 DB 961 AAGACAGCTGGGCGCTGGATACGAGATCCCTATGAGCTTCCATACATGACTACTTTG 1020
 QY 1114 AATATCTTGGACACGATTTTCAAGCTCCACATCAGTCTCTCCATATATGCTAAACCAACA 1173
 DB 1021 AATATCTTGGACACGATTTTCAAGCTCCACATCAGTCTCTCCATATGCTAAACCAACA 1080
 QY 1174 CGATGAGTACCTGAGAGATCAACAGGACCTGTTGAGAACCTTAACTTGTCTGCC 1233
 DB 1081 CTAAAGAGTACCTGAGAGATCAACAGGACCTGTTGAGAACCTTAACTTGTCTGCC 1140
 QY 1234 ACGACCTGGGCGCTCAACAGGAGGATCTCTGAGAGCGCCATCCCTGAGGAGTGAGG 1293
 DB 1141 ATGCCCCCTGGGCGCTCAACAGGAGGATCTCTGAGAGCGCCATCCCTGAGGAGTGAGG 1200
 QY 1294 ATGAGAGCAGAGACGACCTGTGACAGGCACTCTGATCTGCTCTCTGACCAAGATTTG 1353
 DB 1201 ATGAGAGTACAGAGACGACCTGTGACAGGCACTCTGATCTGCTCTCTGATTAAGCATTTG 1260
 QY 1354 CTTGTGAGAGAGATTCTCGATTTCTGAGAGAGAGAGGAGGCGCGCAAGATCTTT 1413
 DB 1261 CTTGTGAGAGAGATTCTCGATTTCTGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1320
 QY 1414 CCAACTTCAAAAAGCCAGAGAGTCAAAAAGAGAGATGAAGAGAGAGAGAGAGAGAGG 1473
 DB 1321 CTAATCTTCAAAAAGCCAGAGAGTCAAAAAGAGAGATGAAGAGAGAGAGAGAGAGG 1380
 QY 1474 AGAAGAAAGAGATCAAGG 1533
 DB 1381 AGAAGAAAGAGATCAAGG 1440
 QY 1534 TCAAGG 1593
 DB 1441 TCAAGG 1499
 QY 1594 CTTGACGTTTCTT 1606
 DB 1500 CTTGACGTTTCTT 1512

RESULT 11
 LOCUS MMU80780 1999 bp mRNA linear ROD 02-SEP-1997
 DEFINITION Mus musculus putative histone deacetylase (HD1) mRNA, partial cds.
 ACCESSION U80780
 VERSION U80780.1 GI:2347179
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 REFERENCE 1 Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Johnson, C.A.
 TITLE Putative histone deacetylase HD1 from M. musculus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1999)
 AUTHORS Johnson, C.A., White, D., O'Neill, L.P. and Turner, B.M.
 TITLE Direct Submission
 JOURNAL Submitted (02-DEC-1996) Department of Anatomy, University of
 Birmingham, Birmingham B15 2TT, U.K.
 COMMENT On Sep 2, 1997 this sequence version replaced gi:1737470.
 FEATURES
 source 1..1999
 /organism="Mus musculus"

/mol_type="mRNA"
 /db_xref="taxon:10090"
 1..1999
 /gene="HD1"
 15..1463
 /gene="HD1"
 /note="similar to human histone deacetylase HD1 and yeast Rpd3"
 /codon_start=1
 /product="putative histone deacetylase"
 /protein_id="AA068398.1"
 /db_xref="GI:2347180"
 /translation="MAQTGTRKRVYDGDVGNYYVGGHPMKPHRIIMTHNLIN
 YGLYRMEIYRPHKANAEMTKYHSDDYIKFLSIRPDNMSEYSKQMPNVEDCPV
 PDGLFEFCOLSTGCVASAVKLNQOTDIAVMNAGLHIAKSEAAGFCYVD1VLAI
 LELIKYHORVLYTIDIDIHGDVEBAFTTDRMTVSPKPYGEYFPGTGLRDIGAK
 GKRYAVNYPRLRSDIDESYEAIPKPYMSKYMEMPQSAVLVOCGSDLSGDRICFNL
 TLGHAKCYEIVKSPNLPMLIMLGGGTYTIRNVAKCTYETNALDTRIEPLEYNDYF
 EYFGPDKLHISPSNMTONTNTEYLEKIKORLEENLRMLPHAVGVOMALPEDAIEE
 SGDEDEDEPDKRISICSDRLIACEEFPSSDESGRGNNSNFKKAKVKTDEKE
 KDEEKEVTEBEKTEBEKPEAKVEEYVLA"
 BASE COUNT 583 a 467 c 480 g 467 t 2 others
 ORIGIN
 Query Match 76.7%; Score 1236.4; DB 10; Length 1999;
 Best Local Similarity 89.0%; Pred. No. 0;
 Matches 1344; Conservative 2; Mismatches 163; Indels 1; Gaps 1;

QY 97 GGGAGCGAGACAGATGCGCGACAGCGAGGCGACCCGAGAGAAAGTCTTACTACTACG 156
 DB 1 GGGCGGGAGACAGATGCGCGACAGCGAGGCGACCAAGAGAAAGTCTTACTACTACG 60
 QY 157 ACGGGAGTGTGAAATTAATTAATTAATGAACAAGGCCCAATGAGCTCCACCGAATCC 216
 DB 61 ACGGGAGTGTGAAATTAATTAATTAATGAACAAGGCCCAATGAGCTCCACCGAATCC 120
 QY 217 GCATGACTCAATATTTGCTGCTCACTATGCTCTTACCGAAAAATGGAATCTATGCC 276
 DB 121 GCATGACTCAATATTTGCTGCTCACTATGCTCTTACCGAAAAATGGAATCTATGCC 180
 QY 277 CTGACAAAGCAATGCTGAGAGATGACCAAGTACCAAGCAATGACTACTAATTAATCT 336
 DB 181 CTGACAAAGCAATGCTGAGAGATGACCAAGTACCAAGCAATGACTACTAATTAATCT 240
 QY 337 TGGGCTCCATCCGTCAGATTAATGATGCGAGTACAGCAAGAGATGAGATTTCAAG 396
 DB 241 TGGGCTCCATCCGTCAGATTAATGATGCGAGTACAGCAAGAGATGAGATTTCAAG 300
 QY 397 TTGATGAGACTGTCCAGTATGATGCGCTGTTGAGTCTGTCAAGTTGCTTACTGTTG 456
 DB 301 TTGATGAGACTGTCCAGTATGATGCGCTGTTGAGTCTGTCAAGTTGCTTACTGTTG 360
 QY 457 GTTCTGTGAGAGTGTCTGTAATTAATTAAGCAGAGAGACATGCGCTGAATTTGG 516
 DB 361 GTTCTGTGAGAGTGTCTGTAATTAATTAAGCAGAGAGACATGCGCTGAATTTGG 420
 QY 517 CTGGGGGCTGACACATGCAAGAGGCGAGGATCTGCTCTGTTAAGTCAATGATA 576
 DB 421 CTGGGGGCTGACACATGCAAGAGGCTGAAGATCTGAAGATCTGCTCTGTTAAGTCAATGATA 480
 QY 577 TGTCTTGTGGCATCTGGAACCTGTAAGTATCAACAGAGAGTCTGTACATTGACATTG 636
 DB 481 TGTCTTGTGGCATCTGGAACCTGTAAGTATCAACAGAGAGTCTGTACATTGACATTG 540
 QY 637 ATATTCACATGATGAGCGGCTGGAAGAGGCTTTCTACACAGAGACCGGGTCAATGCTG 696
 DB 541 ATATTCACATGATGAGCGGCTGGAAGAGGCTTTCTACACAGAGACCGGGTCAATGCTG 600
 QY 697 TGTCTTGTGAATGATGAGATGACTTCCAGAGACTGGGGACCTTACCGGATACCGGGG 756
 DB 601 TGTCTTGTGAATGATGAGATGACTTCCAGAGACTGGGGACCTTACCGGATACCGGGG 660
 QY 757 CTGGCAAGACAGATTTATGTTTAACTACCGGCTCCAGAGCGGAGATTGATGACGAGT 816

```

Db      661 CTGGCAAGGCAAGTACTATGCTGTACATCCACCTGCGAGCGCATTGACGAGAAAT 720
Qy      817 CCTATGAGGCAATTTCAAGCGGCGATGATCCAAAGTAATGAGATGTTCCAGCCTAGTG 876
Db      721 CCTATGAGGCAATTTCAAGCGGCGATGATCCAAAGTAATGAGATGTTCCAGCCTAGTG 780
Qy      877 CGGTGCTTACAGTGTGCTGAGACTCCCTATCTGCGGAGTCGGTATGAGTTCCTCAATC 936
Db      781 CAGTGTCTTACAGTGTGCTGAGACTCCCTATCTGCGGAGTCGGTATGAGTTCCTCAATC 840
Qy      937 TATCTATCAAGGACACGCCAAGTGTGTGAATTTGTCAAGACTTTAACTTCCTATGC 996
Db      841 TGACCAATCAAGGACACGCCAAGTGTGTGAATTTGTCAAGACTTTAACTTCCTATGC 900
Qy      997 TGAATGCGGAGCGGATGTTTACACATTCGTAACTGTCGCCGCTGAGCATATGAGA 1056
Db      901 TGAATGCGGAGCGGATGTTTACACATTCGTAACTGTCGCCGCTGAGCATATGAGA 960
Qy      1057 CAGCTGTGCGCTGATACGAGATCCCTATGAGCTTCCATATGACTATGACTATTTGAAT 1116
Db      961 CAGCGGTGCGCTGATACGAGATCCCTATGAGCTTCCATATGACTATGACTATTTGAAT 1020
Qy      1117 ACTTTGACCAAGATTTCAAGCTTCACATCAGTCTTCCATATGACTATGACTATTTGAAT 1176
Db      1021 ACTTTGACCGGATTTCAAGCTTCACATCAGTCTTCCATATGACTATGACTATTTGAAT 1080
Qy      1177 ATGATGATCTGAGGAAAGATTAACACGAGCTGTTTGAGAACTTAAGTTCCTCCGACG 1236
Db      1081 ACAGATGATCTGAGGAAAGATTAACACGAGCTGTTTGAGAACTTAAGTTCCTCCGACG 1140
Qy      1237 CACCTGCGGCTGCAAGCGAGCGCATCTCTGAGAGCGCATCCCTGAGAGAGTGGCGATG 1296
Db      1141 CCGCTGCGGCTGCAAGCGAGCGCATCTCTGAGAGCGCATCCCTGAGAGAGTGGCGATG 1200
Qy      1297 AGGACGAGACGACCTCTGACAGCGCATCTCGATCTGCTCTGACAAACGAATTCCT 1356
Db      1201 AAGATGAGGAGACCTCTGACAAACGAATCTCGATCTGCTCTGACAAACGAATTCCT 1260
Qy      1357 GTGAGGAAAGTTCCTCGATCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1416
Db      1261 GTGAGGAAAGTTCCTCGATCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Qy      1417 ACTTCAAAAAAGCCAGAGAGTCAAAAAGAGAGATGAAAAAGAGAAAGCCAGAGAGAG 1476
Db      1321 ACTTCAAAAAAGCCAGAGAGTCAAAAAGAGAGATGAAAAAGAGAAAGATCTGAAAGAGA 1380
Qy      1477 AGAAGAGATCACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536
Db      1381 AAAAAAGATCACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy      1537 AGGAGAGAGCCAAAGTGTGCTGATGAGACCTCTGACAGCTGTCGCTTCGTAGTCCCT 1596
Db      1441 AAGAGAGAGCCAAAGTGTGCTG-AGCAAGGTCTGACAGCCCATCTTCTCCCAAGTTCT 1499
Qy      1597 CAGCTTCTT 1606
Db      1500 CACTTCTCCT 1509

```

REFERENCE AUTHORS

- 1 (bases 1 to 252509)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Isman,F.R., Allen,C., Albrooks,S.L., Amaralungu,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benson,J., Bimagne,K., Blankenburg,K., Bonnin,D., Bouck,U., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douchwaite,K.J., Draper,H., Dugan-Rocha,S., Dudin,K.D., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,T., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,U., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,U., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louisged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rofe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshbari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabot,P., Tamerisa,A., Tamerisa,K., Tang,H., Taneey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.
- 2 (bases 1 to 252509)
Unpublished
Worley,K.C.
- 3 (bases 1 to 252509)
Direct Submission
Submitted (23-Oct-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS TITLE JOURNAL

COMMENT

On Nov 13, 2002 this sequence version replaced gi:23195713.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

OY	1349	AATTGCGTGGAGGAAGAAGTTCTCCGATTCGTGAAGAGAGAGAGAGGGGGGCCGCAGAA	1408	
Dd	76399	CATTGCGTGAGGAGGAATTCCTCTGATTCGTGAAGAGAGAGAAAGAGGTCACAGAA	76455	
OY	1409	CTCTTCCAACTTCMAAAAAGCCCAAGACATCMAAACAAGAGATGAAAAAGAAAGACC	1468	
Dd	76459	CTCTTCAACTTCMAAAAAGCCCAAGACATCMAAACAAGAGATGAAAAAGAAAGATCC	76511	
OY	1469	AGAGAGAGAGAGAGAAATCACCCGAAAGAGAGAAAAACCAAGAGAGAGAACCAAGCCAA	1528	
OY	1529	AGGGGTCAAGAGAGAGAGCCCAAGTTGGCTGTAATGAGACTCTCCAGCTCTGGCTTCGCT	1588	
Dd	76579	AGGGGTCAAGAGAGAGAGCCCAAGTTGGCTGTAATGAGACTCTCCAGCTCTGGCTTCGCT	76633	
OY	1589	GAGTCCTCATCAGCTTCTT	1606	
Dd	76638	AAGTCCCTCACTTCCCT	76655	
RESULT 13				
AC096943				
LOCUS				
DEFINITION	AC096943	253360 bp	DNA linear HTG-13-NOV-2002	
ACCESSION	AC096943			
VERSION	AC096943.5			
KEYWORDS	HTG; PHASE1; HTGS_DRAFT; HTGS_ENRICHED.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
REFERENCE				
AUTHORS	1 (bases 1 to 253360)			
	Muzny,D.M., Adams,C., Adio-Otunola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T., Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnan,D., Bowick,J., Bowie,S., Brieva,M., Brown,E., Browm,M., Bryant,N.P., Bunyah,K., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.J., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dedertich,D.A., Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H., Donthaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Galis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez-P., Hodgson,A., Hogues,M., Holloway,C., Hollins,L.E., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Judah,S., Katsison,B., Kelly,S., Khan,U., King,L., Korvah,U., Kovar,C., Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.J., Li,Z., Lichteage,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Locardo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Mattindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Monabatl,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenko,S., Ogun,M., Okumotu,G., Orgunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Qules,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoestlati,N., Sisson,I., Sodergren,E., Sonalkte,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Taylor,C., Tabor,P., Tametris,A., Tameris,K., Tang,H., Tansey,J., Tavak,A., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umami,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,S., Wleceyk,R., Wooden,S., Worley,K.,			

```

Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinert, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 253360)
Morley, K.C.
Direct Submission
Submitted (03-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253360)
Morley, K.C.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23664830.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atl/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: TUNY
Center clone name: CH230-2C19

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 235451 bases at least Q40
Consensus quality: 237249 bases at least Q30
Consensus quality: 238382 bases at least Q20
Estimated insert size: 238893; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html) .
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 248409: contig of 248409 bp in length
* 248410 248509: gap of unknown length
* 248510 249588: contig of 1079 bp in length
* 249589 249688: gap of unknown length
* 249689 251319: contig of 1631 bp in length
* 251320 251419: gap of unknown length
* 251420 253360: contig of 1941 bp in length.
*
Location/Qualifiers
1. 253360
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-2C19"
1. 1149
/note="wgs end_extension
clone_end:T7"
1200..2271
/note="wgs_end_extension

```

```

misc_feature      clone_end:T7"
                  /note=".7695
                  clone_end:T7
                  site:_end:T7
misc_feature      end_sequence:BH262771"
                  /note="wgs_contig"
                  245799..248403
BASE COUNT      73003 a 49394 c 47685 g 68790 t 14488 others
ORIGIN
Query Match      74.6%; Score 1201.8; DB 2; Length 253360;
Best Local Similarity 88.7%; Pred. No. 5.7e-309;
Matches 1348; Conservative 0; Mismatches 162; Indels 9; Gaps 4;
QY 95 CCGGGAGCGAGCAAGATGCGCAGAGCAGGCGACCCGAGGAAAGTCTGTACTCTA 154
DB 75146 CCGGGAGCGAGCAAGATGCGCAGAGCAGGCGACCCGAGGAAAGTCTGTACTCTA 75205
QY 155 CGACGGGAGTGTGAAATTAATTAATTAATGAACAAGCCCAATGAAGCTCAACGAT 214
DB 75206 CGACGGGAGTGTGAAATTAATTAATTAATGAACAAGCCCAATGAAGCTCAACGAT 75265
QY 215 CCGCATGACTATAATTTGCTGCTCAACTATGCTCTTACCGAAATGAATCTATCG 274
DB 75266 CCGCATGACTATAATTTGCTGCTCAACTATGCTCTTACCGAAATGAATCTATCG 75325
QY 275 CCGTCACAAGCCAAATGCTGAGAGATGACCAAGTACACAGCATGACTATTAAT 334
DB 75326 TCCTCACAAGCCAAATGCTGAGAGATGACCAAGTACACAGCATGACTATTAAT 75385
QY 335 CTTCGCTCCATCCGTCAGATTAATGTCGAGTACAGAGAGATGACAGATTCAA 394
DB 75386 CTTCGCTCCATCCGTCAGATTAATGTCGAGTACAGAGAGATGACAGATTCAA 75445
QY 395 CGTTGTGAGAGCTGTCCAGTATTCGATGCTGTTGAGTTCTGTCACTTCTACTCG 454
DB 75446 CGTTGTGAGAGCTGTCCAGTATTCGATGCTGTTGAGTTCTGTCACTTCTACTCG 75505
QY 455 TGGTTCGTGAGAGTGTGAGAACTTAATGAGAGAGAGCAATCGCGGAAATG 514
DB 75506 TGGTTCGTGAGAGTGTGAGAACTTAATGAGAGAGAGCAATCGCGGAAATG 75565
QY 515 GCGTGGGGGCTGACCAATGAGAAAGTCCGAGGATCTGCTTCTGTTACGTAATGA 574
DB 75566 GCGTGGGGGCTGACCAATGAGAAAGTCCGAGGATCTGCTTCTGTTACGTAATGA 75625
QY 575 TATGCTCTTGGCCATCTGGAATGCTTAAGTATCAACAAGGCTGCTTACATTAAT 634
DB 75626 TATGCTCTTGGCCATCTGGAATGCTTAAGTATCAACAAGGCTGCTTACATTAAT 75685
QY 635 TGAATTTACAC-----ATGCTGACGGGTGAGAGAGGCTTCAACCAAGGACCGGGT 688
DB 75686 TGAATTTACACGGGTGAGAGAGGCTTCAACCAAGGACCGGGT 75745
QY 689 CATGACTGTGTCTTCTTAAGTATGAGAGTACTTCCAGAGAACTGGGACCTACAGGGA 748
DB 75746 CATGACTGTGTCTTCTTAAGTATGAGAGTACTTCCAGAGAACTGGGACCTACAGGGA 75805
QY 749 TACCGGGGCTGGCAAGACAAGTATTAATGCTTTAATCAACCGCTCCGAGACGGGATTTGA 808
DB 75806 TATTTGGGGCTGGCAAGACAAGTATTAATGCTTTAATCAACCGCTCCGAGACGGGATTTGA 75865
QY 809 TGAAGAGTCTTAATGAGGCAATTTCAACCGGCTATGTCACCAATTAATGAGATTTCCA 868
DB 75866 TGAAGAGTCTTAATGAGGCAATTTCAACCGGCTATGTCACCAATTAATGAGATTTCCA 75925
QY 869 GCCTAGTGCAGGTGCTTACAGTGTGAGCTCCCTTACTGAGGAGTGGATTGAGTTG 928
DB 75926 GCCTAGTGCAGGTGCTTACAGTGTGAGCTCCCTTACTGAGGAGTGGATTGAGTTG 75985
QY 929 CTTCATCTATCTATCAAGAGACACGCCAAGTGTGTGAATTTGTCAAGACCTTTAACT 988

```

```

DB 75986 CTTCATCTGACCATCAAGACATCCAAAGTGTGAGTCTGTGAAGATTCAACTT 76045
QY 989 GCCTATGCTGATCTGGAGAGCGGTGTTACACCATTTCTGAAGTTGCCCGTGTGAC 1048
DB 76046 GCCTATGCTGATCTGGAGAGAGATGCTTACCATTCANAAATGTGCTGTGTGAC 76105
QY 1049 ATATGAGACAGCTGTGCGCTGTGATACGAGATCCCTTAATGAGCTTCAATGACTA 1108
DB 76106 TTACGAGACAGCTGTGCGCTGTGATACGAGATCCCTTAATGAGCTTCAATGACTA 76165
QY 1109 CTTTGAATTAATTTGAGACCAATTTCAAGTCCCAATAGTCTTCAATTAATGACTA 1168
DB 76166 CTTTGAATTAATTTGAGACCAATTTCAAGTCCCAATAGTCTTCAATTAATGACTA 76225
QY 1169 GAACAGATGATGATCTGTGAGAGATCAACAGCGACTTTTGAACCTTGAATGCT 1228
DB 76226 GAACAGATGATGATCTGTGAGAGATCAACAGCGACTTTTGAACCTTGAATGCT 76285
QY 1229 GCGGCACGACCTGGGGTCCAAACGACGATTCCTGAGGACGCCATCC-CTGAGGAGA 1287
DB 76286 GCGGCACGACCTGGGGTCCAAACGACGATTCCTGAGGACGCCATCCCTGAGGAGA 76345
QY 1288 GTGGCGATGAGAGACGAAGACCTGACGAAGGCGATCTGATCTGCTGACGAAC 1347
DB 76346 GTGGCGATGAGAGACGAAGACCTGACGAAGGCGATCTGATCTGCTGACGAAC 76404
QY 1348 GAATGCTGTGAGAGAGATTCTCCGATTTCTGAAGAGAGAGAGGGGGCGGACAGA 1407
DB 76405 GAATGCTGTGAGAGAGATTCTCCGATTTCTGAAGAGAGAGAGGGGGCGGACAGA 76464
QY 1408 ACTCTTCAACTTCAAAAAAGCCAGAGATCAAAACAGAGATGAAAAAGAAAGACC 1467
DB 76465 ACTCTTCAACTTCAAAAAAGCCAGAGATCAAAACAGAGATGAAAAAGAAAGATC 76524
QY 1468 CAGAGAGAAAGAAAGATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1527
DB 76525 CAGAGAGAAAGAAAGATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 76584
QY 1528 AAGGGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1587
DB 76585 AAGGGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 76643
QY 1588 TGAATCTCTGAGCTTTCTT 1606
DB 76644 CAAGTCCCTCACTTCCCT 76662

```

RESULT 14
 AC125456/c 142347 bp DNA linear ROD 21-DEC-2002
 LOCUS Mus musculus chromosome 18 clone RP24-193W20, complete sequence.
 DEFINITION AC125456
 ACCESSION AC125456
 VERSION AC125456.4 GI:27356751
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 142347)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 142347)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (26-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 REFERENCE 3 (bases 1 to 142347)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-2002) Genome Sequencing Center, 4444 Forest Park
 REFERENCE 4 (bases 1 to 142347)

AUTHORS McPherson, J. D. and Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Dec 21, 2002 this sequence version replaced gi:2606682.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BB0193M20

FEATURES
source
1. 142347
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/clone="RP24-193M20"

BASE COUNT 44878 a 27249 c 26830 g 43390 t
ORIGIN

Query Match 74.2%; Score 1195.6; DB 10; Length 142347;
Best Local Similarity 87.7%; Pred. No. 2.5e-307;
Matches 1328; Conservative 0; Mismatches 184; Indels 2; Gaps 2;

QY 94 GCGGGAGCGGACCAAGATGAGCCAGACGAGGCGACCCGAGAGAAAGTCTTACTACT 153
DB 117641 GCGGGAGCGGACCAAGATGAGCCAGACGAGGCGACCAAGAGAAAGTCTTACTACT 117582
QY 154 ACAGACGGGAGTGTGAAATTAATTATTTATGGAACAAGGCAACCCATGAAGCCCTACCGAA 213
DB 117581 ATGATGGAGAGCTGTGAAATTAATTATTTATGGAACAAGGCGACCCATGAAGCCCTACCGAA 117522
QY 214 TCCGATGATCTCAATATTTGCTGCTCAACTATGCTCTTACCGAAATGGAATCTATC 273
DB 117521 TCTGATGATCTCAATATTTGCTGCTCAACTATGCTCTTACCGAAATGGAATCTATC 117462
QY 274 GCCCTCAAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGCATGATCACTAATTAAAT 333
DB 117461 GTCTCAAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGCATGATCACTAATTAAAT 117402
QY 334 TCTTGGCTCCATCCGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 393
DB 117401 TCTTGGCTCCATCCGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 117342
QY 394 ACCTGGAGGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
DB 117341 ATGTTGGTGAAGCTGTCCGGTATTTATGCTGTTGATGTTCTGTCACTTATCCACCG 117282
QY 454 GTGCTTGTGAGCAAGTGTGAGAACTTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 513
DB 117281 GAGGCTGTGAGCAAGTGTGAGAACTTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 117222
QY 514 GGGCTGGGGGCTGCGACATGCAAGAAAGTCCGAGCATCTGCTTCTTATAGTCAATG 573
DB 117221 GGGCTGGGGGCTGCGACATGCAAGAAAGTCCGAGCATCTGCTTCTTATAGTCAATG 117162
QY 574 ATATCGTCTTGGCCATCTGGAAGTCTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 633
DB 117161 ACATCGTCTTGGCCATCTGGAAGTCTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 117102
QY 634 TTGATATTCACATGCTGAGCGGCTGAGAGAGCGCTTTTACACACGAGACCGGGTCTATGA 693
DB 117101 TTGATATTCACATGCTGAGCGGCTGAGAGAGCGCTTTTACACACGAGACCGGGTCTATGA 117042
QY 694 CTGTGCTCTTCTTAAATATGAG 753
DB 117041 CTGTGCTCTTCTTAAATATGAG 116982
QY 754 GGGCTGGCAAGAGATTAATGCTGTTAATCAACCGCTCCAGAGCGGAGATTGATGACG 813

DB 116981 GGGCTGGCAAGAGATTAATGCTGTTAATCAACCGCTCCAGAGCGGAGATTGATGACG 116922
QY 814 AATCCCTATGAG 873
DB 116921 AATCCCTATGAG 116862
QY 874 GTGCGGTGTCTTATGAGTGTGCTGAGACTCCCTATCTGAGGATGCTGATGAGTGTGCTTCA 933
DB 116861 GTGCGGTGTCTTATGAGTGTGCTGAGACTCCCTATCTGAGGATGCTGATGAGTGTGCTTCA 116802
QY 934 ATCTATCTATCAAGAGACACGCGCAAGTGTGGAATTTGTCAAGACTTTAATCTGCTTCA 993
DB 116801 ATCTGACCATCAAGAGACACGCGCAAGTGTGGAATTTGTCAAGACTTTAATCTGCTTCA 116742
QY 994 TGTGATGTGAG 1053
DB 116741 TGTGATGTGAG 116682
QY 1054 AGACAGCTGTGCGCTGATGATGAGAGATCCCTAATGAGCTTCCATATGATGATGATGATGATGAT 1113
DB 116681 AATGAGCTGTGCGCTGATGATGAGAGATCCCTAATGAGCTTCCATATGATGATGATGATGATGAT 116622
QY 1114 AATGAGCTGTGCGCTGATGATGAGAGATCCCTAATGAGCTTCCATATGATGATGATGATGATGATGAT 1173
DB 116621 AATGAGCTGTGCGCTGATGATGAGAGATCCCTAATGAGCTTCCATATGATGATGATGATGATGATGAT 116562
QY 1174 CGAATGAGATCTGAG 1233
DB 116561 CTAAACAAGTCTGAG 116502
QY 1234 AGCAGCTGTGCGCTGATGATGAGAGATCCCTAATGAGCTTCCATATGATGATGATGATGATGATGAT 1293
DB 116501 ATGCGCTGTGCGCTGATGATGAGAGATCCCTAATGAGCTTCCATATGATGATGATGATGATGATGAT 116442
QY 1294 ATGAG 1353
DB 116441 ATGAG 116382
QY 1354 CCTGTGAG 1413
DB 116381 CCTGTGAG 116322
QY 1414 CCAACTTCAAAAG 1473
DB 116321 CCAACTTCAAAAG 116262
QY 1474 AGAAG 1532
DB 116261 AGAAG 116202
QY 1533 GTCAAG 1592
DB 116201 GTCAAG 116143
QY 1593 CCTGACGTTCTT 1606
DB 116142 CCTGACGTTCTT 116129

RESULT 15
AC129023
LOCUS Mus musculus chromosome UNK clone RP23-430E12, WORKING DRAFT
DEFINITION
ACCESSION AC129023
VERSION AC129023.3 GI:31126887
KEYWORDS HTG: HTGS PHASE1: HTGS DRAFT; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 181964)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL
TITLE
COMMENT

Wilson, R.K.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 181964)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (24-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 181964)
Wilson, R.K.
Direct Submission
Submitted (30-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On May 29, 2003 this sequence version replaced gi:22475688.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0430E12
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-Primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180253 bases at least Q40
Consensus quality: 180732 bases at least Q30
Consensus quality: 181043 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 40509: contig of 40509 bp in length
* 40510 40609: gap of unknown length
* 40610 89491: contig of 48882 bp in length
* 89492 89591: gap of unknown length
* 89592 181964: contig of 92373 bp in length.
Location/Qualifiers
1. 181964
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-430E12"
1. 40509
/note="assembly_name:Contig17"
40610. 89491
/note="assembly_name:Contig18"
89592. 181964
/note="assembly_name:Contig19"
BASE COUNT 57430 a 33929 c 34008 g 56396 t 201 others
ORIGIN

Query Match 74.2% Score 1195.6; DB 2; Length 181964;
Best Local Similarity 87.7%; Pred. No. 2.5e-307;
Matches 1328; Conservative 0; Mismatches 184; Indels 2; Gaps 2;

94 GCGGGAGCGAGCAATGCGCAGACGAGCGGCGAGCAAGCTCTGTTACTACT 153
|||
Db 74900 GCCGGCGGTGAGCAAGATGCGACGACTCAGGGCACAAGAGAAATCTGTTACTACT 74959
|||
Qy 154 ACGAGCGGAGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 213
|||
Db 74960 ATGATGGGAGCGTGGAACTACTATTAATTAATTAATTAATTAATTAATTAAT 75019
|||

Qy 214 TCCGATGACTCATTAATTTGCTCTCACTATGATCTTACCGAAAAATGAATCTATC 273
|||
Db 75020 TCTGATGACTCATCAATTTGCTCTCACTATGATCTTACCGAAAAATGAATCTATC 75079
|||
Qy 274 GCCCTCAAAAGCCAAATGCTGAGAGATGACCAAGTACCAACAGCATGATTAATTAAT 333
|||
Db 75080 GTCTCAAAAGCCAAATGCTGAGAGATGACCAAGTACCAACAGCATGATTAATTAAT 75139
|||
Qy 334 TCTTGGGCTCCATCCGTCGATATACATGCGAGTACAGCAAGATGATGATTAATTAAT 393
|||
Db 75140 TCTTGGGCTCCATCCGTCGATATACATGCGAGTACAGCAAGATGATGATTAATTAAT 75199
|||
Qy 394 ACGTTGGTGAAGACTGTCAGATATTCAGATGAGCTGTTGAGTTCTGATGATCTATCTG 453
|||
Db 75200 ATGTTGGTGAAGACTGTCAGATATTCAGATGAGCTGTTGAGTTCTGATGATCTATCTG 75259
|||
Qy 454 GTGTTCTGTGCGCAAGTGTGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 513
|||
Db 75260 GAGGCTCTGCGCAAGTGTGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75319
|||
Qy 514 GGGCTGGGGGCTGCAACATGCAAAAGATCCGAGGATCTGGCTTCTGTTACGTCATATG 573
|||
Db 75320 GGGCTGGGGGCTGCAACATGCAAAAGATCCGAGGATCTGGCTTCTGTTACGTCATATG 75379
|||
Qy 574 ATATCGTCTTGGCATCTGGAATCTGTAAGTATCACAGAGGCTGTTACATTTGACA 633
|||
Db 75380 ACATCGTCTTGGCATCTGGAATCTGTAAGTATCACAGAGGCTGTTACATTTGACA 75439
|||
Qy 634 TTGATATTCACATGTTGACGCGGTGTAAGAGGCTTTTACACACGAGACGGGTCTATGA 693
|||
Db 75440 TTGATATTCACATGTTGACGCGGTGTAAGAGGCTTTTACACACGAGACGGGTCTATGA 75499
|||
Qy 694 CTGTCCTCTTCAATGATGAGAGTACTCCAGGAACTGGGGGACTTACGGGATACCG 753
|||
Db 75500 CTGTCCTCTTCAATGATGAGAGTACTCCAGGAACTGGGGGACTTACGGGATACCG 75559
|||
Qy 754 GGGCTGCAAAAGCAAGTATTAATGCTGTTAATTAATTAATTAATTAATTAATTAATTAAT 813
|||
Db 75560 GGGCTGCAAAAGCAAGTATTAATGCTGTTAATTAATTAATTAATTAATTAATTAATTAAT 75619
|||
Qy 814 AGTCTATGAGGCGCATTTTAAAGCGGTCATGTCCAAAAGTAAAGAGATTTTCAAGCTTA 873
|||
Db 75620 AATCTATGAGGCGCATTTTAAAGCGGTCATGTCCAAAAGTAAAGAGATTTTCAAGCTTA 75679
|||
Qy 874 GTGCGATGCTTACAGTGTGCTCAGACTCCCTATCTGGGAGATCGGTTAGTTGCTTCA 933
|||
Db 75680 GTGCGATGCTTACAGTGTGCTCAGACTCCCTATCTGGGAGATCGGTTAGTTGCTTCA 75739
|||
Qy 934 ATCTATCTTACAAAGACAGCGCAAGTGTGAAATTTGTCAAGCTTTAACTGCTTA 993
|||
Db 75740 ATCTATCTTACAAAGACAGCGCAAGTGTGAAATTTGTCAAGCTTTAACTGCTTA 75799
|||
Qy 994 TGCATGATCTGAGAGCGGTGTTTACCACTTGTATACGTTGCCGCTGTCGACATATG 1053
|||
Db 75800 TGCATGATCTGAGAGCGGTGTTTACCACTTGTATACGTTGCCGCTGTCGACATATG 75859
|||
Qy 1054 AGACAGCTGTGGCCCTGATACGAGATCCCTATGAGCTTCCATCAATGATGATCTTGG 1113
|||
Db 75860 AGACAGCTGTGGCCCTGATACGAGATCCCTATGAGCTTCCATCAATGATGATCTTGG 75919
|||
Qy 1114 AATCTTGTGACAGATTTTCAAGCTCCATCACTAGTCCCTTCAATGATGATGATGATGATGAT 1173
|||
Db 75920 AATCTTGTGACAGATTTTCAAGCTCCATCACTAGTCCCTTCAATGATGATGATGATGATGAT 75979
|||
Qy 1174 GGAATGATCTGAGAGATCAATCAAGCACTGTTTGAACCTTGAATGCTGCGCC 1233
|||
Db 75980 GGAATGATCTGAGAGATCAATCAAGCACTGTTTGAACCTTGAATGCTGCGCC 76039
|||
Qy 1234 ACGACCTGGGATCCAAAGCAGCGGATCTGAGAGCGGCAATCCGAGAGAGTGGCG 1293
|||
Db 76040 ACGACCTGGGATCCAAAGCAGCGGATCTGAGAGCGGCAATCCGAGAGAGTGGCG 76099
|||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 13:00:53 ; Search time 457 Seconds
(without alignments)
9515.962 Million cell updates/sec

Title: US-09-817-913-2

Perfect score: 1611
Sequence: 1 atgtctgggtctctgccgcg.....tccctcagttcttctccccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N Geneseq_19Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1610.6	100.0	1611	24	ABK87715 Human cDNA encodin
2	1596.6	99.1	22	AAK89554 Human histone deac	
3	1499.6	93.1	2163	22	AAK16188 Human prostate can
4	1499.6	93.1	2378	23	ABV25106 Human prostate exp
5	1498	93.0	2091	24	AAK40884 Human histone deac
6	1498	93.0	2091	24	ABK83674 Human cDNA differe
7	1498	93.0	2111	17	AAK12940 RPD1 transcription
8	1436.2	89.1	1671	25	ABK58329 DNA encoding TAR-d

9	1434.6	89.1	1449	18	KAT86371
10	1433.4	89.0	1674	23	ABK58325
11	1378	85.5	1582	23	AAK81302
12	818	1997	24	AB199512	
13	805.2	50.0	1985	20	AAK32066
14	805.2	50.0	1985	22	AAK90323
15	805.2	50.0	1985	22	AAK89555
16	805.2	50.0	1985	25	ACC50181
17	805.2	50.0	1985	25	ABK34831
18	795.6	49.4	1985	24	ABK87716
19	729	45.3	2145	23	ABL01897
20	640.6	39.8	4755	23	ABL01896
21	502.6	31.2	1943	20	AAK90840
22	501.8	31.1	2019	20	AAK90839
23	501.6	31.1	680	25	ABZ83010
24	499.6	31.0	1990	21	AAZ58259
25	498.2	30.9	1807	22	AAK80350
26	480.2	29.8	1935	25	ABK63632
27	480.2	29.8	1954	22	AAK89556
28	480.2	29.8	1954	24	ABK87717
29	477	29.6	1805	21	AAZ58260
30	460.8	28.6	1576	20	AAK90841
31	459	28.5	1826	20	AAK90837
32	450.2	27.9	455	24	ABV87577
33	434	26.9	1662	23	ABL04121
34	433.4	26.9	1609	21	AAK46043
35	433.4	26.9	1611	21	AAK33022
36	431.8	26.8	1800	22	AAK80351
37	413	25.6	586	22	AAK52097
38	413	25.6	586	22	AAK83742
39	402	25.0	576	22	AAK83742
40	402	25.0	576	22	AAK83742
41	395.4	24.5	680	21	ABO60268
42	392.4	24.4	600	21	AAK16397
43	386.8	24.0	402	22	AAK55632
44	369	22.9	3793	23	ABL04120
45	353.4	21.9	650	24	ABQ60010

ALIGNMENTS

RESULT 1	ID	ABK87715	standard; cDNA; 1611 BP.
XX	XX	ABK87715	
XX	XX	ABK87715	
XX	XX	07-OCT-2002 (first entry)	
DE	XX	Human cDNA encoding Histone deacetylase isoform 1.	
XX	XX	Human; ss; gene; histone deacetylase; HDAC-1; cancer; cytostatic;	
KW	KW	antitumor; tumour suppressor; cell proliferation; tumour;	
KW	KW	programmed cell death; necrotic cell death.	
XX	XX	Homo sapiens.	
OS	OS	Homo sapiens.	
XX	XX	Location/Qualifiers	
FT	FT	key	
FT	FT	CDS	
FT	FT	111..1559	
FT	FT	/*tag= a	
FT	FT	/product= "HDAC-1"	
FT	FT	/transl_except= (pos:750..757,aa:Ile)	
FT	FT	/transl_except= (pos:765..767,aa:Gly)	
FT	FT	/transl_except= (pos:783..788,aa:Tyr)	
FT	FT	/transl_except= (pos:939..941,aa:Thr)	
FT	FT	/transl_except= (pos:1251..1253,aa:Met)	
FT	FT	/transl_except= (pos:1482..1487,aa:Glu+Val)	
FT	FT	/transl_except= (pos:1545..1547,aa:Val)	
XX	XX	US2002061860-A1.	
XX	XX	23-MAY-2002.	

cDNA encoding a hi
DNA encoding HDAC1
DNA encoding novel
Mouse ischaemic co
Human METR2 relate
U31814 cDNA clone.
Human histone deac
Breast cancer asso
Coding sequence SE
Human cDNA encodin
Drosophila melanog
Maize histone deac
Toxicologically re
Rice histone deace
Nucleotide sequenc
Human cDNA #632 di
Human histone deac
Human cDNA encodin
Soybean histone de
Maize histone deac
Maize histone deac
Human colon cancer
Drosophila melanog
Arabidopsis thalia
Arabidopsis thalia
Nucleotide sequenc
Human ovarian PCR-
Human ovarian tumo
Human ovarian PCR-
Human ovarian tumo
Human colon cancer
Human colon cancer
Human breast tumo
Drosophila melanog
Human colon cancer

Query Match	Best Local Similarity	Matches 1611; Conservative	100.0%; Score 1610.6; DB 24; Length 1611; 100.0%; Pred. No. 0;	0; Mismatches	0; Indels	0; Gaps	0;
0Y	1	ATGTTGGGGTCTCTGCCCCGTGCTGCTGTCTCTCCACTCGGTGATCTTGAAGACCA	60				
Db	1	ATGTTGGGGTCTCTGCCCCGTGCTGCTGTCTCTCCACTCGGTGATCTTGAAGACCA	60				
0Y	61	GCTTAGGGRCTCTGTCTCACTCGGGGTAGACCAACGGGGGAGCGAGAAATGGCGCAGA	120				
Db	61	GCTTAGGGRCTCTGTCTCACTCGGGGTAGACCAACGGGGGAGCGAGAAATGGCGCAGA	120				
0Y	121	CGCAGGGGACCCGGAGGAAAGTCTGTTTCTACTACGACGGGGANTTTGGAATTAATTA	180				
Db	121	CGCAGGGGACCCGGAGGAAAGTCTGTTTCTACTACGACGGGGANTTTGGAATTAATTA	180				
0Y	181	ATGACAGAGGCCACCCAAATGAAGCTTCAACCGAATCGCATATCAATTAATTTGCTGCTCA	240				
Db	181	ATGACAGAGGCCACCCAAATGAAGCTTCAACCGAATCGCATATCAATTAATTTGCTGCTCA	240				
0Y	241	ACTATGGTCTCTACCGGAAATATGAAATCTATCGGCCCTCAAAAGCCAAATCTGAGGAGA	300				
Db	241	ACTATGGTCTCTACCGGAAATATGAAATCTATCGGCCCTCAAAAGCCAAATCTGAGGAGA	300				
0Y	301	TGACCAAGTACCAACAGCATGCTCATTTAAATTTTTCGCTCCATCCGTCCAGATACA	360				
Db	301	TGACCAAGTACCAACAGCATGCTCATTTAAATTTTTCGCTCCATCCGTCCAGATACA	360				
0Y	361	TGTGAGAGTACAGCAGAGAGATGACAGATTTCAACGTTGGTGAAGACTGTCCAGATTGCG	420				
Db	361	TGTGAGAGTACAGCAGAGAGATGACAGATTTCAACGTTGGTGAAGACTGTCCAGATTGCG	420				
0Y	421	ATGGCGCTTTGAGATTTCTGTCAAGTTGTCTACGTGTGTCTGTGTGCAAGTGTCTGGAAC	480				
Db	421	ATGGCGCTTTGAGATTTCTGTCAAGTTGTCTACGTGTGTCTGTGTGCAAGTGTCTGGAAC	480				

QY	481	TTATTAAGCAGCAGCAACGCAACATGCGCGTGAATTGGCTGGGGGGCTGACACATGCAAGA	540
Db	481	TTAATTAAGCAGCAGCAGCAACGCAACATGCGCGTGAATTGGCTGGGGGGCTGACACATGCAAGA	540
QY	541	AGTCGAGGAGCATCTGGCTTCTGTTAACGTCAATGATATATCGTCTGGCCATCTGGAACTGC	600
Db	541	AGTCGAGGAGCATCTGGCTTCTGTTAACGTCAATGATATATCGTCTGGCCATCTGGAACTGC	600
QY	601	TAAAGTATCACAGAGGGTGTGTACATTGACATTGATATTACCAATGATGAGCGCGTGG	660
Db	601	TAAAGTATCACAGAGGGTGTGTACATTGACATTGATATTACCAATGATGAGCGCGCTGG	660
QY	661	AAGAGGCGCTTCTACACACGGAACCGGGTCACTGACTGTGCTCTTCAATAGTATGAGAGT	720
Db	661	AAGAGGCGCTTCTACACACGGAACCGGGTCACTGACTGTGCTCTTCAATAGTATGAGAGT	720
QY	721	ACTTCCAGAGACTGGGAGACCTTACGGGATACCGGGGCTGGCAAGACAGATATTATGCTG	780
Db	721	ACTTCCAGAGACTGGGAGACCTTACGGGATACCGGGGCTGGCAAGACAGATATTATGCTG	780
QY	781	TTAATCTACCGGCTCCGAGAGGGGATTGATACGAGTCTCTATAGAGGCGCATTTTCAAGCCGG	840
Db	781	TTAATCTACCGGCTCCGAGAGGGGATTGATACGAGTCTCTATAGAGGCGCATTTTCAAGCCGG	840
QY	841	TCATGTCCAAAGTATGAGATGTGTCCAGGCTAGTGCGGTGTCTTACAGTGTGGCTCAG	900
Db	841	TCATGTCCAAAGTATGAGATGTGTCCAGGCTAGTGCGGTGTCTTACAGTGTGGCTCAG	900
QY	901	ACTGCCCTATCTGGGGATTCGGTTAGGTGTCTTCAATCTATCTATCAAGAGACACGCCAAGT	960
Db	901	ACTGCCCTATCTGGGGATTCGGTTAGGTGTCTTCAATCTATCTATCAAGAGACACGCCAAGT	960
QY	961	GTGTGGAATTTGTCAAGAGCTTTTAACCTGCTATGCTGATGCTGGAGACGGTGTGTACA	1020
Db	961	GTGTGGAATTTGTCAAGAGCTTTTAACCTGCTATGCTGATGCTGGAGACGGTGTGTACA	1020
QY	1021	CCATTCCGTAACGTTGCGCGGTGCTGACATATTGACACAGCTGTGGCCCTTGATATCGGAGA	1080
Db	1021	CCATTCCGTAACGTTGCGCGGTGCTGACATATTGACACAGCTGTGGCCCTTGATATCGGAGA	1080
QY	1081	TCCGTAATGAGCTCCATPACAAATGACTATCTTGAATCTTTGGACAGAGATTTCAAGCTCC	1140
Db	1081	TCCGTAATGAGCTCCATPACAAATGACTATCTTGAATCTTTGGACAGAGATTTCAAGCTCC	1140
QY	1141	ACATCAGTCTCTCCAAATATGACTATACCAAGAACAGATGAGTACCTGAGAGAGATCAAC	1200
Db	1141	ACATCAGTCTCTCCAAATATGACTATACCAAGAACAGATGAGTACCTGAGAGAGATCAAC	1200
QY	1201	AGCGACTGTTTGAACACTTATGAAATGCTGCCGACCGACCTGGGGTCCAAACGAGCGCA	1260
Db	1201	AGCGACTGTTTGAACACTTATGAAATGCTGCCGACCGACCTGGGGTCCAAACGAGCGCA	1260
QY	1261	TTCTCGAGAGACGCCATCTCTGAGAGAGTGGCGATGAGAGACGACGACCTGTACAAGC	1320
Db	1261	TTCTCGAGAGACGCCATCTCTGAGAGAGTGGCGATGAGAGACGACGACCTGTACAAGC	1320
QY	1321	GCATCTCGATCTGCTCTCTGTACCAACGAAATTCCTGTGAGGAAGAAGTCTCCGATTTCTG	1380
Db	1321	GCATCTCGATCTGCTCTCTGTGTACCAACGAAATTCCTGTGAGGAAGAAGTCTCCGATTTCTG	1380
QY	1381	AAGAGGAGGGAGAGAGGGGGCCGCAAGAACTTTCCACTTCAAAAAAGCCAAAGAGATCA	1440
Db	1381	AAGAGGAGGGAGAGAGGGGGCCGCAAGAACTTTCCACTTCAAAAAAGCCAAAGAGATCA	1440
QY	1441	AAACGAGGATGAAAAAGAGAAACCCAGAGAGAAAGAAAGATACCGAGAGAGAGA	1500
Db	1441	AAACGAGGATGAAAAAGAGAAACCCAGAGAGAAAGAAAGATACCGAGAGAGAGA	1500
QY	1501	AAACCAAGAGAGAGAACCCAGAAACCAAAGGGTCTAAGAGAGGCCAAGTTGGCTCGAA	1560
Db	1501	AAACCAAGAGAGAGAACCCAGAAACCAAAGGGTCTAAGAGAGGCCAAGTTGGCTCGAA	1560

Qy	1381	AAGAGGAGGGAGAGGGGGCCGCCAAGAACTCTTCCAACTTTAAAAAACCAGAGAGTCA	1440
Db	1381	AAGAGGAGGGAGAGGGGGCCGCCAAGAACTCTTCCAACTTTAAAAAACCAGAGAGTCA	1440
Qy	1441	AAACAGAGGAGTGAATAAAGAAAGAACCCAGAGAGAGAAAGAAATACCCGAGAGGAGA	1500
Db	1441	AAACAGAGGAGTGAATAAAGAAAGAACCCAGAGAGAGAGAAAGAAATACCCGAGAGGAGA	1500
Qy	1501	AAACCAAGAGAGAGAAAGCCCAAGCCAAAGGGGTCAAGAGAGAGGCCAAAGTTGGCTTGA	1560
Db	1501	AAACCAAGAGAGAGAAAGCCCAAGCCAAAGGGGTCAAGAGAGAGGTCAAGTTGGCTTGA	1560
Qy	1561	TGAGCCTCTCAGGCTCTGGCTTCCGCTGAGGCGCCCTCAGTTCCTTTCCCC	1611
Db	1561	TGAGCCTCTCAGGCTCTGGCTTCCGCTGAGGCGCCCTCAGTTCCTTTCCCC	1611
RESULT 3			
AAFI6188			
ID	AAFI6188 standard; cDNA; 2163 BP.		
XX	AAFI6188;		
XX	13-MAR-2001 (first entry)		
DE	Human prostate cancer antigen nucleotide sequence SEQ ID NO:623.		
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;		
KW	neuroprotective; cytoprotic; cardioactive; immunomodulatory; muscular;		
KW	vulnerable; gastrointestinal; nephrotropic; antiinfective; gynaecological;		
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;		
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;		
KW	wound; infectious disease; ss.		
XX	Homo sapiens.		
XX	MO200055174-A1.		
PD	21-SEP-2000.		
PF	08-MAR-2000; 2000WO-US05988.		
PR	12-MAR-1999; 99US-0124270.		
PA	(HUMA-) HUMAN GENOME SCI INC.		
PA	(ROSE/) ROSEN C A.		
PI	Rosen CA, Ruben SM;		
DR	WPI: 2000-587513/55.		
DR	P-PSDB; AAB56985.		
XX	Prostate cancer associated gene sequences, referred to as prostate		
PT	cancer antigens, useful for treatment, prevention, and diagnosis of		
PT	disorders such as prostate cancer -		
XX	Claim 1; Page 1067-1068; 2338pp; English.		
XX	AAFI5566 to AAFI6505 encode the human prostate cancer associated		
CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.		
CC	The prostate cancer antigens can have neuroprotective, cytoprotic,		
CC	cardioactive, immunomodulatory, muscular, vulnerable, gastrointestinal,		
CC	nephrotropic, antiinfective, gynaecological and antibacterial activities,		
CC	and can be used in gene therapy. The prostate cancer antigen		
CC	polynucleotides may be used for detection of prostate cancer, chromosome		
CC	identification, as chromosome markers, and for numerous other diagnostic		
CC	or research purposes. The prostate cancer antigens may be used to treat		
CC	disorders such as neural, immune, muscular, reproductive,		
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative		
CC	disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to		
CC	AAB57303 represent sequences used in the exemplification of the present		
CC	invention.		

XX	Sequence	2163 BP; 606 A; 490 C; 551 G; 515 T; 1 other;
Query Match	93.1%; Score 1499.6; DB 21; Length 2163;	
Best Local Similarity	99.4%; Pred. No. 0;	
Matches 1505; Conservative	0; Mismatches 9; Indels 0; Gaps 0;	
Qy	97	GGGAGGGGAGCAAAATGCGCGACGCGAGGGCACCCGGAGGAAGTCTGTTACTACTACG 156
Db	71	GGGAGGCGAGCAAAATGCGCGACGCGAGGGCACCCGGAGGAAGTCTGTTACTACTACG 130
Qy	157	ACGGGAGATGTTGAAATTTACTATTATGAGACAGGCCACCCAAATGAGCCTCACCGAATCC 216
Db	131	ACGGGAGATGTTGAAATTTACTATTATGAGACAGGCCACCCAAATGAGCCTCACCGAATCC 190
Qy	217	GCATGACTCATAAATTTGCTGCTCAACTATGCTCTCTACCGAAAATGGAATCTATGGCC 276
Db	191	GCATGACTCATAAATTTGCTGCTCAACTATGCTCTCTACCGAAAATGGAATCTATGGCC 250
Qy	277	CTCAAAAGCCAAATGCTGAGAGATGACAAATGCCAAGGATGACTACATTAAATTC 336
Db	251	CTCAAAAGCCAAATGCTGAGAGATGACAAATGCCAAGGATGACTACATTAAATTC 310
Qy	337	TGCGCTCCATCCGCTCCAGATTAACATGTGCGAGTACAGCAAGCAGATCAGAGATTCAACG 396
Db	311	TGCGCTCCATCCGCTCCAGATTAACATGTGCGAGTACAGCAAGCAGATCAGAGATTCAACG 370
Qy	397	TTGGTGAAGACTGTCCAGATTTCCAGTGGCCCTGTTGAGTTCGTCACTGTCTTCACTGGTG 456
Db	371	TTGGTGAAGACTGTCCAGATTTCCAGTGGCCCTGTTGAGTTCGTCACTGTCTTCACTGGTG 430
Qy	457	GTTCTGTGAGCAAGTGTCTGTAACTTAAATGACGACGACGAGATGCGCGTGAATTGGG 516
Db	431	GTTCTGTGAGCAAGTGTCTGTAACTTAAATGACGACGAGATGCGCGTGAATTGGG 490
Qy	517	CTGGGGGGCTGACACATGCAAAAGATCCGAGGATCGGCTTCTGTTACGTCATGATA 576
Db	491	CTGGGGGGCTGACACATGCAAAAGATCCGAGGATCGGCTTCTGTTACGTCATGATA 550
Qy	577	TGCTCTTGCGCCATCTGGAACTGCTAAAGTATCACAGAGGTGCTGTACATTGACATYG 636
Db	551	TGCTCTTGCGCCATCTGGAACTGCTAAAGTATCACAGAGGTGCTGTACATTGACATYG 610
Qy	637	ATATTACCATGTGTGACGGCGGTGGAAGAGCGCTTCTACACACGAGCCGGGTATGACTG 696
Db	611	ATATTACCATGTGTGACGGCGGTGGAAGAGCGCTTCTACACACGAGCCGGGTATGACTG 670
Qy	697	TGTCCTTTCATTAAGTATGAGAGTACTTCCAGGAACGAGGAGCTTACATGAGATGACGAGG 756
Db	671	TGTCCTTTCATTAAGTATGAGAGTACTTCCAGGAACGAGGAGCTTACATGAGATGACGAGG 730
Qy	757	CTGGCAAAAGCAAAATTTATGCTGTTAACACCGCTCCGAGACGGGATGATGATGACGAGT 816
Db	731	CTGGCAAAAGCAAAATTTATGCTGTTAACACCGCTCCGAGACGGGATGATGATGACGAGT 790
Qy	817	CCATAGAGCCATTTTCAAGCCGGTCAATGTCCAAAGTATGAGATGTTCCAGCTTATGTG 876
Db	791	CCATAGAGCCATTTTCAAGCCGGTCAATGTCCAAAGTATGAGATGTTCCAGCTTATGTG 850
Qy	877	CGGTGTGCTTACATGTGTGCTCAACATCTCCATCTGGGGATCGGTTAGGTTGCTTCAATC 936
Db	851	CGGTGTGCTTACATGTGTGCTCAACATCTCCATCTGGGGATCGGTTAGGTTGCTTCAATC 910
Qy	937	TATCATCAAAAGACAGCCCAAGTGTGGAATTTGCAAGAGCTTTAACTGCTCTATGC 996
Db	911	TATCATCAAAAGACAGCCCAAGTGTGGAATTTGCAAGAGCTTTAACTGCTCTATGC 970
Qy	997	TGATGTGAGAGGCGGTGTATCACCATGTGTAAGTTCGCCGGTGTGACATATGAGA 1056
Db	971	TGATGTGAGAGGCGGTGTATCACCATGTGTAAGTTCGCCGGTGTGACATATGAGA 1030
Qy	1057	CAGCTGTGGCCCTTGATACGAGATCCCTTAATGAGCTTCCATACATGACTACTTTGAT 1116

Db 1031 CAGCTGTGCGCCTGATGAGGAGATCCCTAATGAGCTTCCTCAATGACTACTTGTGAAT 1090
Qy 1117 ACTTTGACCAAGATTTCAAGCTCAGATGCTCTTCAATATGATCAACAGAACAGCA 1176
Db 1091 ACTTTGACCAAGATTTCAAGCTCAGATGCTCTTCAATATGATCAACAGAACAGCA 1150
Qy 1177 ATGAGTACTGTGAGAGATCAACAGAGATGTTTGAAGAACTTTAGAACTGTCGCCAGC 1236
Db 1151 ATGAGTACTGTGAGAGATCAACAGAGATGTTTGAAGAACTTTAGAACTGTCGCCAGC 1210
Qy 1237 CACCTGGGGTCCCAAGCAGAGGATTCCTGAGAGACCCCTCTCTGAGAGAGAGTGGGATG 1296
Db 1211 CACCTGGGGTCCCAAGCAGAGGATTCCTGAGAGACCCCTCTCTGAGAGAGAGTGGGATG 1270
Qy 1297 AGGACGAGAGAGAGAGAGAGATGCTGATGCTGCTCTCTGAGAGAGAGTGGGATG 1356
Db 1271 AGGACGAGAGAGAGAGAGATGCTGATGCTGCTCTCTGAGAGAGAGTGGGATG 1330
Qy 1357 GTGAGAGAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1416
Db 1331 GTGAGAGAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1390
Qy 1417 ACTTCAAAAAAGCAGAGAGATGATCAACAGAGATGATCAACAGAGAGATGATCAAC 1476
Db 1391 ACTTCAAAAAAGCAGAGAGATGATCAACAGAGATGATCAACAGAGAGATGATCAAC 1450
Qy 1477 AGAAGAGATGATCAACAGAGAGATGATCAACAGAGAGATGATCAACAGAGAGATGAT 1536
Db 1451 AGAAGAGATGATCAACAGAGAGATGATCAACAGAGAGATGATCAACAGAGAGATGAT 1510
Qy 1537 AGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
Db 1511 AGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1570
Qy 1597 CACGTTCTTTCC 1610
Db 1571 CACGTTCTTTCC 1584

RESULT 4

ABV25106
ID ABV25106 standard; cDNA; 2378 BP.

AC ABV25106;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 25097.

XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX 17-FEB-2000; 2000US-183319P.

XX 16-MAR-2000; 2000US-189862P.

XX 25-MAY-2000; 2000US-207454P.

XX 09-JUN-2000; 2000US-211314P.

XX 18-JUL-2000; 2000US-219007P.

XX 13-DEC-2000; 2000US-255281P.

PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
PS Claim 1, Page 4885; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (1) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.
SQ Sequence 2378 BP; 659 A; 559 C; 619 G; 541 T; 0 other;
Query Match 93.1%; Score 1499.6; DB 23; Length 2378;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 97 GGGAGGCGAGAGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156
Db 76 GGGAGGCGAGAGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 135
Qy 157 AGGGAGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 216
Db 136 AGGGAGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 195
Qy 217 GCATGACTATATTTGCTGCTCAACTATGCTCTACCGAAAAATGAAATCTATCGCC 276
Db 196 GCATGACTATATTTGCTGCTCAACTATGCTCTACCGAAAAATGAAATCTATCGCC 255
Qy 277 CTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGCATGATCAATTAATTTCT 336
Db 256 CTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGCATGATCAATTAATTTCT 315
Qy 337 TGGCTCCATCCGCTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 396
Db 316 TGGCTCCATCCGCTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 375
Qy 397 TTGATGAGAGAGTGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 456
Db 376 TTGATGAGAGAGTGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 435
Qy 457 GTTCTGTGCAAGTGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 516
Db 436 GTTCTGTGCAAGTGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 495
Qy 517 CTGGGGGCTGACCAATGCAAGAGAGTCCGAGGATCTGGCTTCTGTTACGTCATAGATA 576
Db 496 CTGGGGGCTGACCAATGCAAGAGAGTCCGAGGATCTGGCTTCTGTTACGTCATAGATA 555
Qy 577 TCGTCTTGCCATCCGCAATGCTAAAGTATCACAGAGAGGAGTGTACATTAATTAATTA 636
Db 556 TCGTCTTGCCATCCGCAATGCTAAAGTATCACAGAGAGGAGTGTACATTAATTAATTAAT 615
Qy 637 AATTCACCATGATGAGCGGCTGGAAGAGAGCTTTACACACGAGCGGGATCATAGATG 696
Db 616 AATTCACCATGATGAGCGGCTGGAAGAGAGCTTTACACACGAGCGGGATCATAGATG 675
Qy 697 TGTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
Db 676 TGTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 735
Qy 757 CTGGCAAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JR;
XX WPI; 2001-662795/76.
XX

```

Db      736 CTGGCAAGGCAATATATATCTGTAACTACCGCTCCGAGAGGGGATGATGACAGT
Qy      817 CATTATAGGSCATTTTAAAGCCGCTCATATCCAAAGTAATGAGATGTCAGCCTATG
Db      796 CTTATAGGSCATTTTAAAGCCGCTCATATCCAAAGTAATGAGATGTCAGCCTATG
Qy      877 CGGTGATCTTACAGTGTGCTCAGACTCCCTATCTGGGAGATCGGTAGTGTCTTCAATC
Db      856 CGGTGATCTTACAGTGTGCTCAGACTCCCTATCTGGGAGATCGGTAGTGTCTTCAATC
Qy      937 TATCTATCAAGGACAGCCCAAGTGTGTGAATTTGTCAAGAGCTTTAACTGCTATGC
Db      916 TAACTATCAAGGACAGCCCAAGTGTGTGAATTTGTCAAGAGCTTTAACTGCTATGC
Qy      997 TGATGCTGGAGGGGGGTGTACACATTCGTAAAGTTCGCGGTGTGTGACATATAGA
Db      976 TGATGCTGGAGGGGGGTGTACACATTCGTAAAGTTCGCGGTGTGTGACATATAGA
Qy      1057 CAGCTGTGGCCCTGATACGAGATCCCTAATGAGCTTCCATATGACTTACTTTGAAT
Db      1036 CAGCTGTGGCCCTGATACGAGATCCCTAATGAGCTTCCATATGACTTACTTTGAAT
Qy      1117 ACTTTGAGCAGATTTCAAGCTCAATCAGTCTTCCATATGACTTACTTCAAGACAGA
Db      1096 ACTTTGAGCAGATTTCAAGCTCAATCAGTCTTCCATATGACTTACTTCAAGACAGA
Qy      1177 ATGAGTACCTGGAGAAATCAAAACAGGACCTGTTGAAACCTTGAATGTGCGGACG
Db      1156 ATGAGTACCTGGAGAAATCAAAACAGGACCTGTTGAAACCTTGAATGTGCGGACG
Qy      1237 CACCTGGGGTCAAAACGAGCGCATCTCTGAGACGCGCATCCCTGAGAGAGTGGCATG
Db      1216 CACCTGGGGTCAAAACGAGCGCATCTCTGAGACGCGCATCCCTGAGAGAGTGGCATG
Qy      1297 AGAGCAGAGCAGACCTTGACAAAGGCGCATCTGATCTGCTCTGACAAACGAATGCTT
Db      1276 AGAGCAGAGCAGACCTTGACAAAGGCGCATCTGATCTGCTCTGACAAACGAATGCTT
Qy      1357 GTGAGGAAAGTCTCCGATCTGAAAGAGGAGGAGGGGGCGGAAAGACCTTTCCA
Db      1336 GTGAGGAAAGTCTCCGATCTGAAAGAGGAGGAGGGGGCGGAAAGACCTTTCCA
Qy      1417 ACTTCAAAAAAGCAGAGAGTCAAAAACAGAGATGAAAAAGAAAGACCCAGAGAGA
Db      1396 ACTTCAAAAAAGCAGAGAGTCAAAAACAGAGATGAAAAAGAAAGACCCAGAGAGA
Qy      1477 AGAAAGAAATCAGCGAAAGAGAGAAAAACAAAGAGAGAGCCAGAAAGGAGTCA
Db      1456 AGAAAGAAATCAGCGAAAGAGAGAAAAACAAAGAGAGAGCCAGAAAGGAGTCA
Qy      1537 AGAGAGAGCCCAATTTGGCCCTGAATGACCTTCAGCTGGGCTTCTGCTGAGTCCCT
Db      1516 AGAGAGAGCCCAATTTGGCCCTGAATGACCTTCAGCTGGGCTTCTGCTGAGTCCCT
Qy      1597 CACGTTCTTCTCC 1610
Db      1576 CACGTTCTTCTCC 1589

```

```

XX      OS Homo sapiens.
XX      FH Key
XX      FT CDS
XX      FT /tag= a
XX      FT /product= "Human HDA-1 protein"
XX      MO200250244-A2.
XX      PD 27-JUN-2002.
XX      PE 07-DEC-2001; 2001MO-US46518.
XX      PR 19-DEC-2000; 2000US-0745167.
XX      PA (ISIS-) ISIS PHARM INC.
XX      PI Monia BP, Wyatt JR;
XX      DR WPI; 2002-519880/55.
XX      DR P-PDB; AAD40884.
XX      PT Antisense compounds targeted against polynucleotides encoding Histone
XX      PT deacetylase 1 useful for treating hyperproliferative conditions, e.g.
XX      PT cancer of hematopoietic, lymphoid, myeloid or breast, or a viral
XX      PT infection -
XX      PS Example 13; Page 99-102; 120pp; English.
XX      CC The present invention relates to antisense compounds, compositions and
XX      CC methods for modulating the expression of Histone deacetylase 1 (HDAL).
XX      CC Sequences of the invention are useful for inhibiting the expression of
XX      CC HDAL in cells or tissues and for treating an animal having a disease or
XX      CC condition associated with HDAL e.g., hyperproliferative condition, which
XX      CC is cancer of hematopoietic, lymphoid, myeloid or breast or a condition
XX      CC resulting from a viral infection. Antisense compounds either alone or in
XX      CC combination with other antisense compounds or therapeutics can be used
XX      CC as tools in differential and/or combinatorial analyses to elucidate the
XX      CC expression patterns of a portion or the entire complement of genes
XX      CC expressed within cells and tissues. They are commonly used as research
XX      CC reagents and diagnostics. They may also be useful prophylactically such
XX      CC as to prevent or delay infection, inflammation or tumour formation. The
XX      CC present DNA sequence is human HDA-1 DNA.
XX      SQ Sequence 2091 BP; 552 A; 485 C; 544 G; 510 T; 0 other;
XX      Query Match 93.0%; Score 1498; DB 24; Length 2091;
XX      Best Local Similarity 99.3%; Pred. No. 0;
XX      Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy      97 GGGAGGCGAGCAAGATGGCGCAGACGAGGGCACCAGGAGAAAGTCTGTTACTACTACG 156
Db      50 GGGAGGCGAGCAAGATGGCGCAGACGAGGGCACCAGGAGAAAGTCTGTTACTACTACG 109
Qy      157 ACGGGAGTGTGGAATTAATTAATGACCAAGGCCACCCCAATGAAGCCTACCGAATCC 216
Db      110 ACGGGAGTGTGGAATTAATTAATGACCAAGGCCACCCCAATGAAGCCTACCGAATCC 169
Qy      217 GCATGACTAATTTGCTGCTCACTATGCTCTTAACGAAAAATGAAATCTATTCGCC 276
Db      170 GCATGACTAATTTGCTGCTCACTATGCTCTTAACGAAAAATGAAATCTATTCGCC 229
Qy      277 CTCGAAAGCCATGCTGAGAGAGATGACCAAGTACCAAGCGATGACTATTAATTTCT 336
Db      230 CTCGAAAGCCATGCTGAGAGAGATGACCAAGTACCAAGCGATGACTATTAATTTCT 289
Qy      337 TGGGCTCCATCGTCCAGATTAACATGTCGAGTACAGCAGAGATGACAGATTCACAG 396
Db      290 TGGGCTCCATCGTCCAGATTAACATGTCGAGTACAGCAGAGATGACAGATTCACAG 349
Qy      397 TTGGTGAAGACTGTCCAGTATTCGATGGCTGTTTGAAGTTCTGAGTTGTCTACTGCTG 456

```

```

RESULT 5
AAD40884
ID AAD40884 standard; DNA; 2091 BP.
AC AAD40884;
XX
XX 30-OCT-2002 (first entry)
XX
XX Human histone deacetylase 1 DNA.
XX
XX Human, histone deacetylase 1; HDAL; enzyme; hyperproliferative condition;
XX viral infection; prophylactic; inflammation; phosphothioate backbone;
XX tumour; antisense; cytostatic; virucide; gene; ds.

```


Db	350	TTGGTGGAGACATGCCAGTATTCATGGAGCCCTGTTGAGTTCTGTCAGTGTCTACCTGATG	409
Qy	457	GTTCTGTGGCAAGTGTCTGTAAACTTAATAGCAGCAGACGACATGGCGGTGAATTTGGG	516
Db	410	GTTCTGTGGCAAGTGTCTGTAACTTAATAGCAGCAGACGACATGGCTGTGAATTTGGG	469
Qy	517	CTGGGGGCGTGCACATGCAAAAGAAAGTCCAGGCGATCGGGCTTCGTATGATCAATGTA	576
Db	470	CTGGGGGCGTGCACATGCAAAAGAAAGTCCAGGCGATCTGGCTTCTGTATGATCAATGTA	529
Qy	577	TGCTCTTGGCCATCCTGGAACTGCTAAAGTATCAACAGAGGTCCTGTACATTTGACATTTG	636
Db	530	TGCTCTTGGCCATCCTGGAACTGCTAAAGTATCAACAGAGGTCCTGTACATTTGACATTTG	589
Qy	637	ATATTTACCATGTGTGACGGCGTGAAGAGGCTTCTACCAACGACCGGATCATGACTG	696
Db	590	ATATTTACCATGTGTGACGGCGTGAAGAGGCTTCTACCAACGACCGGATCATGACTG	649
Qy	697	TGTCCTTTCATAAATATGAGAAAGTACTTCCAGGAACGTGGGACCTACCGGATACCGGGG	756
Db	650	TGTCCTTTCATAAAGTATGAGAAAGTACTTCCAGGAACGTGGGACCTACCGGATATCGGGG	709
Qy	757	CTGGCAAAAGCAAGTATTAATCTGTAACTACCCGCTCCGAGAGGGGATTTGATGACAGT	816
Db	710	CTGGCAAAAGCAAGTATTAATCTGTAACTACCCGCTCCGAGAGGGGATTTGATGACAGT	769
Qy	817	CTATATGAGGCCATTTTCAACGCGGTATGTCCAAAGTATGAGAGATTTCCAGCCTATG	876
Db	770	CTATATGAGGCCATTTTCAACGCGGTATGTCCAAAGTATGAGAGATTTCCAGCCTATG	829
Qy	877	CGGATGTCTTAACAATGTGTGCTCAACATCCCATCTGAGGAGTCGGTTAGTTAGTTCAATC	936
Db	830	CGGATGTCTTAACAATGTGTGCTCAACATCCCATCTGAGGAGTCGGTTAGTTAGTTCAATC	889
Qy	937	TATCTATCAAAAGACACGCGCAAGTGTGGAATTTGTCTCAAGAGCTTTAACCTGCTATGC	996
Db	890	TATCTATCAAAAGACACGCGCAAGTGTGGAATTTGTCTCAAGAGCTTTAACCTGCTATGC	949
Qy	997	TGATGTGGAGGCGGTGTTCACAATTCGTMAAGTTGCCGGTGTGACATATGAGA	1056
Db	950	TGATGTGGAGGCGGTGTTCACAATTCGTMAAGTTGCCGGTGTGACATATGAGA	1009
Qy	1057	CAGCTGTGGCCCTGTGATTCGAGATTCCTTAATGAGCTTCCATACATGACTAATTGAT	1116
Db	1010	CAGCTGTGGCCCTGTGATTCGAGATTCCTTAATGAGCTTCCATACATGACTAATTGAT	1069
Qy	1117	ACTTTGGACCAAGTATTCAGAGTCCACATCACTCCCTCCAAATATGACTATACAGAAACAGA	1176
Db	1070	ACTTTGGACCAAGTATTCAGAGTCCACATCACTCCCTCCAAATATGACTATACAGAAACAGA	1129
Qy	1177	ATGAGTACTGTGAGAAAGTCAAAACAGGCACTGTTTGGAGACCTTAAGATGTGCGCGACG	1236
Db	1130	ATGAGTACTGTGAGAAAGTCAAAACAGGCACTGTTTGGAGACCTTAAGATGTGCGCGACG	1189
Qy	1237	CACCTGGGGTCAAAACGAGGCGATTCCTGTGAGACGCCATCCTGAGGAGGTGGCGATG	1296
Db	1190	CACCTGGGGTCAAAATGACGCGATTCCTGTGAGACGCCATCCTGAGGAGGTGGCGATG	1249
Qy	1297	AGGACGAAGACGACCCCTGACAAAGGCGATCTGATCTGTCTCTGACAAAGAAATTTGCT	1356
Db	1250	AGGACGAAGACGACCCCTGACAAAGGCGATCTGATCTGTCTCTGACAAAGAAATTTGCT	1309
Qy	1357	GTCAGGAAGATTTCTCGAATCTGAAGAGAGGAGAGGGGGGCGCGCAAGAACTCTTCCA	1416
Db	1310	GTCAGGAAGATTTCTCGAATCTGAAGAGAGGAGAGGGGGGCGCGCAAGAACTCTTCCA	1369
Qy	1417	ACTTCAAAAAAGCCAAAGAGTCAAAACAGAGATGAAAAAGAAAGAACCCCAAGAGAGA	1476
Db	1370	ACTTCAAAAAAGCCAAAGAGTCAAAACAGAGATGAAAAAGAAAGAACCCCAAGAGAGA	1429
Qy	1477	AGAAAGGAATACCGAAGAGAGAAACCAAGAGAGAGAACCCGAAGCCAAAGGGGTCA	1536
Db	1430	AGAAAGGAATACCGAAGAGAGAAACCAAGAGAGAGAACCCGAAGCCAAAGGGGTCA	1489

Oy		1537	AGAGGAGGGCAGTTGGCTGAATGCACCTTCACCTCGCTTCGTAGTCCCT	1596
Dd		1490	AGAAGAGAGGTCAATTGGCTGAATGACCTTCCTCACCTCGCTTCGTAGTCCCT	1549
Oy		1597	CACGTTTTCTTTCCC	1610
Dd		1550	CACGTTTCTTCCCC	1563
Db				
RESULT 6				
ID	ABK83674			
XX	ABK83674	standard; cDNA; 2091 BP.		
AC				
XX	ABK83674;			
DT	14-AUG-2002	(first entry)		
XX				
DE				
XX				
KW		Human CDNA differentially expressed in granulocytic cells #245.		
KM		Human; ss; granulocytic cell; DNA chip; bacterial infection;		
KM		viral infection; parasitic infection; protozoal infection;		
KM		fungal infection; sterile inflammatory disease; psoriasis;		
KM		rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;		
KM		cardiac reperfusion injury; renal reperfusion injury; AIDS;		
KM		adult respiratory distress syndrome; inflammatory bowel disease;		
KM		Crohn's disease; ulcerative colitis; periodontal disease;		
KM		granulocyte activation; chronic inflammation; allergy.		
OS		Homo sapiens.		
XX				
PN		WO200228999-A2.		
PD		11-APR-2002.		
PF		03-OCT-2001; 2001WO-US30821.		
PR		03-OCT-2000; 2000US-237189P.		
PA		(GENE-) GENE LOGIC INC.		
Pt		Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;		
DR		WPI; 2002-435328/46.		
PT		Detecting granulocyte activation by detecting differential expression		
PT		of genes associated with granulocyte activation, which serves as		
PT		diagnostic markers that is useful for monitoring disease states and		
PT		drug toxicity -		
PS		Claim 1; SEQ ID No 245; 114pp; English.		
XX				
XX		The invention relates to detecting (M1) granulocyte (GC) activation		
CC		(GCA), by detecting the level of expression of gene(s) (Gs) identified by		
CC		DNA chip analysis as given in the specification, and comparing		
CC		the expression level to an expression level in an unactivated		
CC		Gc, where differential expression of Gs is indicative of GCA.		
CC		Also included are modulating (M2) GA by contacting GC with an agent		
CC		that alters the expression of at least one gene in Gs; (2) screening (M3)		
CC		for an agent capable of modulating GCA or an inflammation (especially		
CC		chronic) in a tissue, an allergic response in a subject, exposure of a		
CC		subject to a pathogen or sterile inflammatory disease using the		
CC		gene expression profile; (3) detecting (M4) an inflammation (especially		
CC		chronic) in a tissue, an allergic response in a subject, exposure of a		
CC		subject to a pathogen or sterile inflammatory disease, by detecting the		
CC		level of expression in a sample of the tissue of gene(s) from Gs, where		
CC		the level of expression of the gene is indicative of inflammation;		
CC		(4) treating (M5) an inflammation (especially chronic) or in a tissue,		
CC		an allergic response in a subject, exposure of a subject to a pathogen		
CC		or sterile inflammatory disease, by contacting a tissue having		
CC		inflammation with an agent that modulates the expression of gene(s)		
CC		from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for		
CC		modulating Ga; M3 is useful for screening an agent capable of modulating		

FT 5'UTR 1..63
 FT /+tag= b
 FT 1513..2111
 FT /+tag= c
 EP08112-A1.
 XX
 PD 24-APR-1996.
 XX
 PF 21-SEP-1995; 95SEP-0114884.
 XX
 PR 22-SEP-1994; 94UP-0227876.
 XX
 PA (CANC-) CANCER INST.
 PA (EISA) EISAI CO LTD.
 XX
 PI Furukawa Y, Nakamura Y;
 XX
 DR MPI; 1996-202003/21.
 DR P-PSDB; AAR88919.
 XX
 PT New human transcriptional control protein RPD1 - used to develop
 PT prods. for study and diagnosis involving the protein, partic. for
 PT gene analysis
 PS
 PS Claim 3; Page 10-12; 15pp; English.
 CC AA12940 encodes a transcriptional control protein, RPD1, derived from
 CC a human foetal lung cDNA library. The gene encoding RPD1 is localised
 CC at 1p34.1 on the short arm of chromosome 1, this is a region where a
 CC deletion is recognised in mammary and gastric carcinomas. The RPD1
 CC gene is an important gene and has been found to be expressed in all
 CC the following human tissues: heart, kidney, liver, lung, pancreas,
 CC placenta, skeletal muscle, large intestine, peripheral leukocyte,
 CC ovary, prostate, small intestine, spleen, testis and thymus gland;
 CC but is not present in brain tissue. RPD1 cDNA exhibits homology with
 CC the yeast transcription factor RPD3 and is believed to have a similar
 CC function. RPD1 and DNA encoding it can be used in studying the
 CC interactions of the protein with DNA and other proteins in the cell.
 CC
 XX
 XX Sequence 2111 BP; 572 A; 485 C; 544 G; 510 T; 0 other;
 SQ
 Query Match 93.0%; Score 1498; DB 17; Length 2111;
 Best local similarity 99.3%; Pred. No. 0;
 Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 97 GGGAGCGAGCAAGATGCGCAGACGCGGAGGAAAGTCTGTACTACTAG 156
 DB 50 GGGAGCGAGCAAGATGCGCAGACGCGGAGGAAAGTCTGTACTACTAG 109
 QY 157 ACGGGAGTGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 216
 DB 110 ACGGGAGTGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 169
 QY 217 GCATGACTCAATATTGCTGCTCACTATGCTCTACACGAAATTAATTAATTAAT 276
 DB 170 GCATGACTCAATATTGCTGCTCACTATGCTCTACACGAAATTAATTAATTAAT 229
 QY 277 CTCACAAAGCAATGCTGAGAGATGACCAAGTACCAAGCATGACTCAATTAATTC 336
 DB 230 CTCACAAAGCAATGCTGAGAGATGACCAAGTACCAAGCATGACTCAATTAATTC 289
 QY 337 TGGGCTTCATCCGTCAGATTAATGTCGAGTACAGCAAGCATGAGTCAAGTCAAG 386
 DB 290 TGGGCTTCATCCGTCAGATTAATGTCGAGTACAGCAAGCATGAGTCAAGTCAAG 349
 QY 397 TTGGTGAAGACTGTCAGATTAATGTCGAGTACAGCAAGCATGAGTCAAGTCAAG 456
 DB 350 TTGGTGAAGACTGTCAGATTAATGTCGAGTACAGCAAGCATGAGTCAAGTCAAG 409
 QY 457 GTTCTGTGCAAGTGTGTGAACTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 516
 DB 410 GTTCTGTGCAAGTGTGTGAACTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 469

QY 517 CTGGGGGCTTGACCATGCAAGAAAGTCCGAGGCACTTGTCTTGAATGATGATA 576
 DB 470 CTGGGGGCTTGACCATGCAAGAAAGTCCGAGGCACTTGTCTTGAATGATGATA 529
 QY 577 TCGTCTTGCCCATCTCGAAGTCTGTAAGTATCACAAGAGGCTGCTGATGATGATG 636
 DB 530 TCGTCTTGCCCATCTCGAAGTCTGTAAGTATCACAAGAGGCTGCTGATGATGATG 589
 QY 637 AATATCAACATGATGACCGGCTGGAAGAGGCTTCTGACCAAGAGCGGGATGATGATG 696
 DB 590 AATATCAACATGATGACCGGCTGGAAGAGGCTTCTGACCAAGAGCGGGATGATGATG 649
 QY 697 TGTCTTTTCAATGATGAGAGAGTACTTCCAGAACTGGGAGACTGAGGATACCGGGAG 756
 DB 650 TGTCTTTTCAATGATGAGAGAGTACTTCCAGAACTGGGAGACTGAGGATACCGGGAG 709
 QY 757 CTGGCAAGACAAAGTATTAATGCTGTTAACTACCCGCTCCGAGACGGGATTTGATGACGAGT 816
 DB 710 CTGGCAAGACAAAGTATTAATGCTGTTAACTACCCGCTCCGAGACGGGATTTGATGACGAGT 769
 QY 817 CCTATGAGGCAATTTTCAAGCCGGTCAATGCTCAAGTATGAGAGATTTCCAGCCATGAG 876
 DB 770 CCTATGAGGCAATTTTCAAGCCGGTCAATGCTCAAGTATGAGAGATTTCCAGCCATGAG 829
 QY 877 CGATGCTTTACAGTGTGCTCAGACTCCCTATCTGGGAGTGGTTAGGTTGCTTCAATC 936
 DB 830 CGATGCTTTACAGTGTGCTCAGACTCCCTATCTGGGAGTGGTTAGGTTGCTTCAATC 889
 QY 937 TATCTATCAAGAGACACGCAAGTGTGAGAAATTTGTCAAGAGCTTTAACTCTGATG 996
 DB 890 TATCTATCAAGAGACACGCAAGTGTGAGAAATTTGTCAAGAGCTTTAACTCTGATG 949
 QY 997 TGAATGTGAGAGCGGTGTGTTACCATTCGTAAGCTTCCGCTGCTGACATATGAGA 1056
 DB 950 TGAATGTGAGAGCGGTGTGTTACCATTCGTAAGCTTCCGCTGCTGACATATGAGA 1009
 QY 1057 CAGCTGTGACCTGATGAGAGATCCCTATGAGAGCTTCAATGATGATGATGAT 1116
 DB 1010 CAGCTGTGACCTGATGAGAGATCCCTATGAGAGCTTCAATGATGATGATGATGAT 1069
 QY 1117 ACTTTGACCAAGATTTCAAGCTCCATCATGATGCTTTCAATATGACTAAACGAACGAG 1176
 DB 1070 ACTTTGACCAAGATTTCAAGCTCCATCATGATGCTTTCAATATGACTAAACGAACGAG 1129
 QY 1177 ATGAGTACTGGAAGAGATCAACAGCAGCTGTTTGAAGACCTTGAAGTCTGCGCAG 1236
 DB 1130 ATGAGTACTGGAAGAGATCAACAGCAGCTGTTTGAAGACCTTGAAGTCTGCGCAG 1189
 QY 1237 CACCTGGGGTCCAAACGAGGGGATTCCTGAGGAGGCTCCCTGAGAGAGTGGGATG 1296
 DB 1190 CACCTGGGGTCCAAACGAGGGGATTCCTGAGGAGGCTCCCTGAGAGAGTGGGATG 1249
 QY 1297 AGGACGAAGACGACCTCTGACCAAGCGCATGCTGATCTCTGACCAAGATTTGCT 1356
 DB 1250 AGGACGAAGACGACCTCTGACCAAGCGCATGCTGATCTCTGACCAAGATTTGCT 1309
 QY 1357 GTGAGGAAGAGTCTTCCGATTTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1416
 DB 1310 GTGAGGAAGAGTCTTCCGATTTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
 QY 1417 ACTTCAAAAAGCGAAGAGTCAAAACAGAGGATGAAAAGGAAAGCCCGAGAGAG 1476
 DB 1370 ACTTCAAAAAGCGAAGAGTCAAAACAGAGGATGAAAAGGAAAGCCCGAGAGAG 1429
 QY 1477 AGAAGAGATCAACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536
 DB 1430 AGAAGAGATCAACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489
 QY 1537 AGAAG 1596
 DB 1490 AGAAG 1549


```

Dh 781 TGTGGCTCAGACTCCCTATCTGGGGATCGTTAGTTGCTTCAATCTAATCAAAAGA 840
Qy 951 CAGGCCAAGTGTGTAATTTGTCAAGAGCTTTAAGCTGCTATGATGCTGGAGGC 1010
Db 841 CAGCCCAAGTGTGTAATTTGTCAAGAGCTTTAAGCTGCTATGATGCTGGAGGC 900
Qy 1011 GGTGGTTACACATTCCTTAAAGTTCGGGCTGCTGAGCATATGAGACAGCTGGCCCTG 1070
Db 901 GGTGGTTACACATTCCTTAAAGTTCGGGCTGCTGAGCATATGAGACAGCTGGCCCTG 960
Qy 1071 GATACGAGATCCCTAATAGCTTCCATACATGATCTTGAATCTTTGAGCCAGAT 1130
Db 961 GATACGAGATCCCTAATAGCTTCCATACATGATCTTGAATCTTTGAGCCAGAT 1020
Qy 1131 TTCAAGCTCCACATGCTCTTCCATATGATCTTAAACGAAACAGAAATGAGTACCTGAG 1190
Db 1021 TTCAAGCTCCACATGCTCTTCCATATGATCTTAAACGAAACAGAAATGAGTACCTGAG 1080
Qy 1191 AAGATCAAAACAGCACTGTTGAGAACCTTAAAGTGTGCGACGACCTGGGGTCCAA 1250
Db 1081 AAGATCAAAACAGCACTGTTGAGAACCTTAAAGTGTGCGACGACCTGGGGTCCAA 1140
Qy 1251 ACGGAGCGATTCCTGAGAGCGCAATCCCTGAGAGAGTGGCGATGAGAGCAAGACGAC 1310
Db 1141 ATGAGAGCGATTCCTGAGAGCGCAATCCCTGAGAGAGTGGCGATGAGAGCAAGACGAC 1200
Qy 1311 CCGTACAGGCGATCTGATCTGCTCTGACAAAGAAATGCTGAGAGAGATTC 1370
Db 1201 CCGTACAGGCGATCTGATCTGCTCTGACAAAGAAATGCTGAGAGAGATTC 1260
Qy 1371 TCCGATTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430
Db 1261 TCCGATTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1330
Qy 1431 AAGAGAGTCAAAACAGAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTACC 1490
Db 1321 AAGAGAGTCAAAACAGAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTACC 1380
Qy 1491 GAAAGAGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1550
Db 1381 GAAAGAGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1551 TTGGCTCTGA 1559
Db 1441 TTGGCTCTGA 1449

```

```

XX 19-APR-2002; 2002MO-KR00730.
XX 20-APR-2001; 2001KR-0021449.
PR 18-APR-2002; 2002KR-0021307.
XX (HURM/) HUR M.
PI HUR M, Chong DL;
XX WPI; 2003-093103/08.
DR P-PSDB; ABG72558.
XX
PT New fusion proteins, useful for repressing HIV transcription regulating
PT expression of AIDS viral RNA to inhibit the proliferation of virus and
PT production of resistant virus
XX
PS Claim 5; Page 49-51; 60pp; English.
XX
CC This invention relates to a novel fusion protein which may be used to
CC repress human immunodeficiency virus (HIV) transcription. The protein
CC comprises a transcription inhibitory polypeptide or its compound and a
CC polypeptide or its compound which recognizes the RNA strand around
CC expression control regions or viral long terminal repeat (LTR) promoter
CC cis-acting elements. The fusion proteins of the invention may have Anti
CC -HIV activity and may be used as an inhibitor of HIV Tat. The fusion
CC proteins of the invention are useful for repressing HIV transcription
CC regulating expression of AIDS viral RNA to inhibit the proliferation of
CC virus and production of resistant virus. The method of repressing HIV
CC transcription is useful for treating AIDS. The present sequence
CC represents the DNA sequence encoding the HDAC1-Tat fusion protein of
CC the invention.
SQ Sequence 1674 BP; 472 A; 385 C; 444 G; 373 T; 0 other;
Qy 111 ATGGCGGAGCGAGGCGACCCGAGGAAAGTGTGTTACTATACAGCGGGAGTGTGGA 170
Db 1 ATGGCGGAGCGAGGCGACCCGAGGAAAGTGTGTTACTATACAGCGGGAGTGTGGA 60
Qy 171 AATTACTATATGAGCAAGGCGACCCATGAAAGCTTCAACCGAATCCGATGATCATAT 230
Db 61 AATTACTATATGAGCAAGGCGACCCATGAAAGCTTCAACCGAATCCGATGATCATAT 120
Qy 231 TTGCTGCTCAATGATGCTCTTACCGAAAGTGAATCTATGCGCTCAAAAGCCAT 290
Db 121 TTGCTGCTCAATGATGCTCTTACCGAAAGTGAATCTATGCGCTCAAAAGCCAT 180
Qy 291 GCTGAGAGATGACCAAGTACCAAGAGATGATCACTTAATTTTGGCGCTCATCCGT 350
Db 181 GCTGAGAGATGACCAAGTACCAAGAGATGATCACTTAATTTTGGCGCTCATCCGT 240
Qy 351 CCAATTAACATGTCGAGATCAGACAGATGATCAAGATTTCAACGTTGAGAGATGT 410
Db 241 CCAATTAACATGTCGAGATCAGACAGATGATCAAGATTTCAACGTTGAGAGATGT 300
Qy 411 CCAATTAACATGTCGAGATCAGACAGATGATCAAGATTTCAACGTTGAGAGATGT 470
Db 301 CCAATTAACATGTCGAGATCAGACAGATGATCAAGATTTCAACGTTGAGAGATGT 360
Qy 471 GCTGTAAACTTAATTAAGACGAGACGAGATGCGCTGAATTTGGTGGGGCGCTGAC 530
Db 361 GCTGTAAACTTAATTAAGACGAGACGAGATGCGCTGAATTTGGTGGGGCGCTGAC 420
Qy 531 CATGCAAGAAGTCCGAGGATCGGCTTGTGTTAGCTCATGATATGCTTGGCCATC 590
Db 421 CATGCAAGAAGTCCGAGGATCGGCTTGTGTTAGCTCATGATATGCTTGGCCATC 480
Qy 591 CTGGAATGCTAAAGTATCAACAGAGGCTGCTATGATGATGATATTCATCATGATGT 650

```

```

RESULT 10
ABS58325
ID ABS58325 standard; DNA; 1674 BP.
XX
AC ABS58325;
XX
DT 04-MAR-2003 (first entry)
XX
DE DNA encoding HDAC1-TAT dmt fusion protein.
XX
KW HIV; human immunodeficiency virus; Tat; HIV Tat inhibitor; virus;
KW HIV transcription; AIDS; acquired immunodeficiency syndrome;
KW HDAC1-TAT dmt; gene; de; human.
XX
OS Chimeric - Human immunodeficiency virus type 1.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1674
FT /tag=a
FT /product="HDAC1-TAT fusion protein"
XX
XX MO200285948-A1.
XX
XX 31-OCT-2002.

```



```

Db      481 CTGGAACCTGTAAGATACACGAGGGGCTGTACATTGACATTAATTCACCATGGT 540
Qy      651 GACGGGTGGAGAGGCTCTCTACACCAAGGCGGTCTCATGCTGTCTCTTCAATAG 710
Db      541 GACGGGTGGAGAGGCTCTCTACACCAAGGCGGTCTCATGCTGTCTCTTCAATAG 600
Qy      711 TATGAGAGTACTTCCAGAGAACTGGGACCTACGAGATACCGGGCTGGCAAAACAAG 770
Db      601 TATGAGAGTACTTCCAGAGAACTGGGACCTACGAGATACCGGGCTGGCAAAACAAG 660
Qy      771 TATTAAGCTGTAACTACCGGCTCCGAGAGGGAATTGATGACAGAGTCTATGAGGCATT 830
Db      661 TATTAAGCTGTAACTACCGGCTCCGAGAGGGAATTGATGACAGAGTCTATGAGGCATT 720
Qy      831 TTCAGCCCGGTCACTGTCCAAAGTAATGAGATGTTCCAGGCTGAGTGGTCTTACAG 890
Db      721 TTCAGCCCGGTCACTGTCCAAAGTAATGAGATGTTCCAGGCTGAGTGGTCTTACAG 780
Qy      891 TGTGGCTCAGACTCCCTATCTGGGATCGGTTAGGTTCTCAATCTATCTATCAAGA 950
Db      781 TGTGGCTCAGACTCCCTATCTGGGATCGGTTAGGTTCTCAATCTATCTATCAAGA 840
Qy      951 CAGGCCAAGTGTGGAATTTGTCAGAGCTTTAAGCTGCTATGCTGATGCTGGAGGC 1010
Db      841 CAGGCCAAGTGTGGAATTTGTCAGAGCTTTAAGCTGCTATGCTGATGCTGGAGGC 900
Qy      1011 GGTGGTTACACATTCGTAACTGTCGCGGTGGGATGATGAGACAGTGTGGCCCTG 1070
Db      901 GGTGGTTACACATTCGTAACTGTCGCGGTGGGATGATGAGACAGTGTGGCCCTG 960
Qy      1071 GATACGAGATCCCTATGATGATGCTTCCATACATGACTACTTGAATATCTTGGACGAG 1130
Db      961 GATACGAGATCCCTATGATGATGCTTCCATACATGACTACTTGAATATCTTGGACGAG 1020
Qy      1131 TTCAAGCTCAGATCAGTCTCTCAATATGACTTAACCAAGACGAAATGATGCTGGAG 1190
Db      1021 TTCAAGCTCAGATCAGTCTCTCAATATGACTTAACCAAGACGAAATGATGCTGGAG 1080
Qy      1191 AAGATCAACAGGACGCTTGTGAACCTTAAGATGCTGCGGACGACCTGGGGTCCAA 1250
Db      1081 AAGATCAACAGGACGCTTGTGAACCTTAAGATGCTGCGGACGACCTGGGGTCCAA 1140
Qy      1251 ACCGAGCGCATCTCTGAGAGAGCCATCCCTGAGAGAGTGGCGATGAGACGAAAGCAG 1310
Db      1141 ATGCAAGCGCATCTCTGAGAGAGCCATCCCTGAGAGAGTGGCGATGAGACGAAAGCAG 1200
Qy      1311 CCTGACAGCGCATCTCTGATCTCTCTCTGACAAAGAAATGCTGTGAGAAAGTTTC 1370
Db      1201 CCTGACAGCGCATCTCTGATCTCTCTCTGACAAAGAAATGCTGTGAGAAAGTTTC 1260
Qy      1371 TCCGATTCTGAAGAAGAGAGAGGGGGCGCAAGAACTTCCAACTTCAAAAAAGCC 1430
Db      1261 TCCGATTCTGAAGAAGAGAGAGGGGGCGCAAGAACTTCCAACTTCAAAAAAGCC 1320
Qy      1431 AAGAGAGTCAAAACAGAGATGAAAAAGAAAAAGACCCAGAGGAGAAAGAAATCACC 1490
Db      1321 AAGAGAGTCAAAACAGAGATGAAAAAGAAAAAGACCCAGAGGAGAAAGAAATCACC 1380
Qy      1491 GAAAGAGAGAAAAACAAGAGAGAGAAAGCCAAAGGGGTCAAGAGAGAGCCCAAG 1550
Db      1381 GAAAGAGAGAAAAACAAGAGAGAGAAAGCCAAAGGGGTCAAGAGAGAGCCCAAG 1440
Qy      1551 TTGGCTGGAAT 1561
Db      1441 TTGGCTGGAAT 1451

```

```

RESULT 11
AAS81302
ID AAS81302 standard; cDNA; 1582 BP.
XX AC AAS81302;
XX

```

```

DT      13-FEB-2002 (first entry)
XX      DNA encoding novel human diagnostic protein #17106.
DE      Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX      food supplement; medical imaging; diagnostic; genetic disorder; ss.
KM      Homo sapiens.
XX      MO200175067-A2.
PN      11-OCT-2001.
PD      30-MAR-2001; 2001WO-US08631.
PF      31-MAR-2000; 2000US-0540217.
PR      23-AUG-2000; 2000US-0649167.
XX      (HYSB-) HYSBQ INC.
PA      Drmanac RT, Liu C, Tang YT;
PI      WPI; 2001-639362/73.
XX      P-PSDB; ABG17115.
DR      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensic, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity.
XX      Claim 1; SEQ ID No 17106; 103bp; English.
PS      The invention relates to isolated polynucleotide (I) and
XX      polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC      polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC      and gene mapping, and in recombinant production of (II). The
CC      polynucleotides are also used in diagnostics as expressed sequence tags
CC      for identifying expressed genes. (I) is useful in gene therapy techniques
CC      to restore normal activity of (II) or to treat disease states involving
CC      (II). (II) is useful for generating antibodies against it, detecting or
CC      quantitating a polypeptide in tissue, as molecular weight markers and as
CC      a food supplement. (II) and its binding partners are useful in medical
CC      imaging of sites expressing (II). (I) and (II) are useful for treating
CC      disorders involving aberrant protein expression or biological activity.
CC      The polypeptide and polynucleotide sequences have applications in
CC      diagnostics, forensics, gene mapping, identification of mutations
CC      responsible for genetic disorders or other traits to assess biodiversity
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. AAS64197-AAS94564 represent novel human
CC      diagnostic coding sequences of the invention.
CC      Note: The sequence data for this patent did not appear in the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pat_sequences.
XX      Sequence 1582 BP; 422 A; 376 C; 441 G; 343 T; 0 other;
SQ

```

```

Query Match      85.5%; Score 1378; DB 23; Length 1582;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1475; Conservative 0; Mismatches 40; Indels 13; Gaps 6;

```

```

Qy      97 GGGAGCGAGCAAGATGCGCAGACGACGAGGACCCCGAGAGAAAGTCTTACTACTACG 156
Db      55 GGGAGCGAGCAAGATGCGCAGACGACGAGGACCCCGAGAGAAAGTCTTACTACTACG 114
Qy      157 ACGGGAGTGTGAAATTAATTAATGATGAC-AAAGCCACCAATGAGGCTC-ACCGAAT 214
Db      115 ACGGGAGTGTGAAATTAATTAATGAGGAGAGGAGCCCAATGAAAGCTCGACCGAAT 174
Qy      215 CCGCATGACTCAATTAATTTGCTGCTCAATATGCTCTCAACGAAAAATGAAATCATATG 274
Db      175 CCGCATGACTCAATTAATTTGCTGCTCAATATGCTCTCAACGAAAAATGAAATCATATG 234
Qy      275 CCTCAAAAGCAATGTGAGAGATGACCAAGTACCAAGAGATGACTATTAATT 334

```


CC angiogenesis both in vitro and in vivo. They can be used for treating
CC cancer and other disorders related to angiogenesis including abnormal
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC endometrial bleeding disorders, diabetic retinopathy, some forms of
CC macula degeneration, haemangiomas, and arterial-venous malformations.
CC They may be useful in treating deficiencies or disorders of the immune
CC system, by activating or inhibiting the proliferation, differentiation,
CC or mobilisation (chemotaxis) of immune cells. The etiology of these
CC immune deficiencies or disorders may be genetic, somatic, such as
CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC toxins), or infectious. They can also be used to treat inflammatory
CC conditions, both chronic and acute conditions. The products can also be
CC used for detection and diagnosis. AA232002 to AA232080, and AA49503 to
CC AA49511 represent sequences given in the exemplification of the present
CC invention.

XX Sequence 1985 BP, 626 A, 360 C, 454 G, 545 T, 0 other;

Query Match 50.0%; Score 805.2; DB 20; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3.5e-234;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

```
QY 126 GGCACCCGAGAGAAAGTCTGTTACTACTACGACGGGATGTTGGAATTACTATTATGGA 185
DB 223 GCGCGCAAAAAAAGTCTGCTACTACTACGACGGTATTTGGAAATATTATTATGGA 282
QY 186 CAAAGCCCAACCAATGAAGCTCAACGAAATCGCAATGACTCAATTTGCTGCTCAACTAT 245
DB 283 CAGGCTCAATCCCATGAAAGCTCATGAAATCGCAATGACCAATTAATTTGCTTTAAATAT 342
QY 246 GGTCTTACCAAAAAAATGGAATCTATCGCCCTCAAAAGCCAAATGCTGAGAGATGACC 305
DB 343 GCGTTATACAGAAAAAATGGAATATATAGCCCAATAAAGCACTGCCGAAAGATGCA 402
QY 306 AAGTACCAAGGATGACTATTAATTTCTGGGCTCCATCCCTCGAATATCATGTGCG 365
DB 403 AAAATATCAAGTATGATATATATCAATTTCTACGCTCAATTAACCGAATTAACATGCT 462
QY 366 GAGTACGCAAGGATGATGAGATTCAAAGTGGTGAAGCTGCGCATTTGATGATGCG 425
DB 463 GAGTATGATGAGATGATGATATTTATGATGAGAAATGTCGACGCTTGAATGGA 522
QY 426 CTGTTGAGTTCTGCACTGTTCTACTGCTGTTCTGCAAGTGTGTTGAACTTAT 485
DB 523 CTCTTTGAGTTTGTGCACTCTCAACTGCGGTTCACTGCTGAGGCTGTAAGTTAAC 582
QY 486 AAGCAGCAGACGACATGCGCGTGAATGGCTGGGGCTGACCATGCAAAAGATCC 545
DB 583 CGAACACAGACTGATATGCTGTTAATGGCTGGAGATTAACATGCTTAAAGAAATAC 642
QY 546 GAGGATCTGCTCTGTTAGTCAATGATTCGTTGGCCATCCCTGGAATCTGTAAG 605
DB 643 GAAAGATCAGGATCTGTTACGTTAATGATATGTTGCTGCAATCTTGAATTAAG 702
QY 606 TATCAGCAGAGGATGCTGTACATTTGATTAATTCACATGCTGAGCGGCTGGAAG 665
DB 703 TATCATGAGAGAGCTTATATATTTGATTAATTAATTAATGATGATGTTGAAGAA 762
QY 666 GCGTTTCAACACGAGACCGGCTGATGCTGTCTTCAATTAAGTATGAGAGTACTTC 725
DB 763 GCTTTTATATCAACAGATCGTGATATGACGCTATCACTCCATTAATATGGGAAATACCTT 822
QY 726 CCAGGAAGTGGGACCTACGAGATTAACGGGGCTGGCAAAAGCAATTTATGCTGTTAAC 785
DB 823 CTGACACAGAGAGCTTGAAGGATTTGCTCTGAAAAGCAAAATCTATGCTGTCAT 882
QY 786 TACCGCTCCGAGACGGGATTTGATGACGAGTCTTATGAGGCAATTTTCAAGCCGATCAG 845
DB 883 TTTCCATATGTTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 942
QY 846 TCCTAAAGTATGAGATGTTCCAGCTATGTCGGTGTCTTACAGTGTGCTCAGACTCC 905
DB 943 TCAAGGATGATGAGATGATGATCAACCTAGTGTGATTAACAGTGTGCTGAGACTCA 1002
```

```
QY 906 CTATCTGGGATGCGTTAGTGTGCTTCAATCTATCTATCAAGGACACGCGAATGTGTG 965
DB 1003 TTATCTGATGATAGCTGCGTTGTTTCAATCTATCAAGTCAAAAGGCTACGTTAAATGTGTA 1062
QY 966 GAATTTGTCAAGACTTTAACTGCTATGCTGATGCTGAGAGCGGTGTTACACCAT 1025
DB 1063 GAAGTTTAAAACTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1122
QY 1026 CGTACGTTGCGGCTGCTGACATATGAGACAGCTGTGGCTCTGATAGAGAGATCCCT 1085
DB 1123 CGTAAATGCTGATGTTGAGACATATGAGACAGCTGTGGCTCTGATAGAGAGATCCCT 1182
QY 1086 AATAGCTTCATCAATGATGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1145
DB 1183 AATAGCTTCATCAATGATGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1242
QY 1146 AGTCTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205
DB 1243 AGTCTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
QY 1206 CTGTTTGAAGACCTTGAATGCTGTGCGGACGACCTTGGGTCCAAAGCGAGGATTCCT 1265
DB 1303 TTGTTGAAAATTTGGCATGTTAATCTGATGACCTGTGTGTCAGATGCAAGCTTATTC 1362
QY 1266 GAGGACGCAATCCCTGAGAGAGTGGCGATGAGAGACGAAAGCAACCTGACAAAGCGCATC 1325
DB 1363 GAGGATGCTTTTATGAGAGACAGTGAAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 1422
QY 1326 TCGATCTGCTCTCTGACAAACGAATTCCTGTGAGAGAGATTCCTGATTTGAAAG 1385
DB 1423 TCTATTCGACATCAAGCAAGCGGATGCTTGTGATGAGAAATTTCCAGATTTGAGAT 1482
QY 1386 GAGGAGAGGAGGCGCGCAAGAACTTCCACTTCAAAAAACCAAGAGATGCAAAACA 1445
DB 1483 GAAAGAGAGAGAGAGTGAAGAAATGTGCTGATCATTAAGAAAGACAAAGAAAGCTGGA 1542
QY 1446 GAGGATGAAAAGAGAAAGACCCAGAGAGAGAGAAAGAAATCAACCGAAGAGAGAAACC 1505
DB 1543 ATTGAAGAGATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTA 1602
QY 1506 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1551
DB 1603 TCCAAAGACAAAGTGTGTAAGAAAGAGATTAACCAAGAAAGCAAAAT 1648

RESULT 14
AAC90323
ID AAC90323 standard; DNA; 1985 BP.
XX
AC AAC90323;
XX
DT 19-MAR-2001 (first entry)
XX
DE U1814 cDNA clone.
XX
XX
XX METH: metalloproteinase; thrombospondin; angiogenesis inhibition;
XX cancer therapy; benign tumour; ocular angiogenic disease;
XX rheumatoid arthritis; psoriasis; wound healing; endometriosis;
XX vasculogenesis; granululation; hypertrophic scar; nonunion fracture;
XX scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;
XX coronary collateral; cerebral collateral; arteriovenous malformation;
XX ischaemic limb angiogenesis; Osler-Weber syndrome; wound granulation;
XX plaque neovascularisation; telangiectasia; haemophilic joint; EST;
XX angiodioma; fibromuscular dysplasia; expressed sequence tag;
XX Crohn's disease; atherosclerosis; birth control; ss.
XX
XX Unidentified.
XX
XX OS
XX
XX PN MO200071577-A1.
XX
XX PD 30-NOV-2000.
XX
```

PE 25-MAY-2000; 2000MO-US14462.
XX 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.
PR 22-FEB-2000; 2000US-0183792.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (SMIX-) SMITHKLINE BEECHAM CORP.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (TRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
PA (JONA/) JONAK Z L.
PA (TRUL/) TRULLI S H.
PA (FORN/) FORNWALD J A.
PA (TERR/) TERRETT J A.
XX
PI IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH,
PI Fornwald JA, Terrett JA;
XX
XX WPI: 2001-025136/03.
XX
PT METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT inhibit angiogenesis in the treatment of disorders such as cancer,
XX rheumatoid arthritis and psoriasis -
XX
PS Claim 14; Pages 745-746; 768pp; English.
XX
XX The present invention relates to human METH1 and METH2, (ME for
XX metalloproteinase and TH for thrombospondin; see AAB50002 and AAB50003).
XX The present sequence is an expressed sequence tag (EST) for METH. METH
XX can be used for inhibiting angiogenesis in an individual, and for
XX treating cancer, benign tumours, an ocular angiogenic disease,
XX rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
XX vasculogenesis, granulomatous, hypertrophic scars, nonunion fractures,
XX scleroderma, trichoma, vascular adhesions, myocardial angiogenesis,
XX coronary collateralis, cerebral collateralis, arteriovenous malformations,
XX ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
XX neovascularisation, telangiectasia, haemophilic joints, angiodiroma,
XX fibromuscular dysplasia, wound granulation, Crohn's disease or
XX atherosclerosis. METH can also be used in birth control. METH can also
XX be used in diagnostic methods for the prognosis of cancer.
SQ Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;
Query Match 50.0%; Score 805.2; DB 22; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3.5e-234;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
QY 126 GGCACCCGAGGAAAGTCTGTTACTTACGACGGGATGTTGAAATTACTATTATGGA 185
DB 223 GGGCGGCAAAAAAAGTCTGTTACTTACGACGGGATGTTGAAATTACTATTATGGA 282
QY 186 CAGGCCACCAATGAAAGCTCCACCGAATCGGATGACTATATTTGCTGCTCAATAT 245
DB 283 CAGGGTATCCATGAAGGCTCATAGATCCGATGACCAATTAATTGCTGTTAAATAT 342
QY 246 GGTCTTACCAAAATGAAATCTATGCTCCACCAAGCAATGCTGAGATGACC 305
DB 343 GGGCTATACGAAATGAAATATATAGGCCCATTAAGCCATGCGCAAGAAATACA 402
QY 306 AAGTACCAAGCGATGACTTAAATTCTTGCGCTCCATCCGCTCAATTAACATGTC 365
DB 403 AATATATACAGTGTATGATATCAAAATTTCAAGTCAATTAAGACCAAGTAAATGCT 462
QY 366 GAGTACGACGAGATGACGAGATTCAAGTTGGTGGAGAGCTGTCAGATTGAGATGCG 425
DB 463 GAGTATGATGAGAGATGACGATATATTTAAATGTTGGAAGAAATGTTCCAGGTTGATGGA 522
QY 426 CTGTTGAGTTCTGTCAAGTTGCTACTGTGTTCTGTGCGAAGTGTGAAACTTAAT 485

DB 523 CTCTTGAAGTTTGTGAGCTCTCACTGCGGCTTCAAGTTGCTGAGCTGTGAAGTTAAAC 582
QY 486 AAGCAGACAGCGACATCCGCGTAATTTGGGCTGGGGGCTCGACCATGCAAGAAAGTCC 545
DB 583 CGACACAGACTATATGCTGTAAATTTGGGCTGGAGATTACATATGCTTAAGAAATAC 642
QY 546 GAGGATCTGGCTTCTGTATACGTCAATGATATGCTTGGCCATCTCTGAACTGCTAAG 605
DB 643 GAAGCATCAGGATTTCTGTATCGTAAATGATATGCTGCTGCTGCTGCTGCTGCTAAG 702
QY 606 TATCACAGAGGCTGCTGATCAATTGACATTTATTCACATGCTGACGCGCTGGAAGAG 665
DB 703 TATCATCAGAGACTTATATATGATTAATGATTAATCATCATGATGATGCTGTGAAGAA 762
QY 666 GCTTCTACACCAAGGACCGGCTGACGCTGCTCTTCAATAGTATGAGAGTCTTC 725
DB 763 GCTTTTATACAAAGATCGTAAATGACGTAATCCATTAATATGAGAAATCTTT 822
QY 726 CCAGAACTGGGACCTACCGGATACCGGGCTGGCAAGACAAATTAATGCTGTTAAC 785
DB 823 CTGCGACAGGACATTTGAGGATATTTGGTCTGGAAGAAAGCAATATCTATGCTCAAT 882
QY 786 TACCCGCTCCGAGACGGGATTTGATGACAGTCTATGAGCCATTTTCAAGCCGTCATG 845
DB 883 TTTCCATGCTGATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 942
QY 846 TCCAAATGATGAGAGATGCTCCAGCTGAGCGGCTGCTTAAAGTGTGCTGACCTCC 905
DB 943 TCAAGGTGATGAGAGATGATTAACCTAGTGTGCTGATTTAAGTGTGATGACCTCA 1002
QY 906 CTATCTGGGATGCTGATGCTGCTTCAATCTATCTATCAAGACACGCAAGTGTG 965
DB 1003 TTATCTGCTGATGATGATGCTGCTTCAATCTATCAAGTCAAGTCAAGTCAAGTCAAGT 1062
QY 966 GAAATTTGCAAGCTTTTAACTGCTATGCTGATGCTGAGGCGGTGTTACACCAT 1025
DB 1063 GAAGTTGAAAACCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1122
QY 1026 CTTAAGCTTCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1085
DB 1123 CGTAAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1182
QY 1086 AATGAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1145
DB 1183 AATGAGTGCATATATATGATATCTTGTGATATTTTGGACCAAGCTTCAATGCTATTT 1242
QY 1146 AGTCTTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1205
DB 1243 AGTCTTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1302
QY 1206 CTGTTTGAAGACTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265
DB 1303 TTGTTTGAAGAAATTTGCGATGTTTAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1362
QY 1266 GAGGAGCCATCCCTGAGGAGGTGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1325
DB 1363 GAGGATGCTGTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1422
QY 1326 TCGATCTGCTCTGATCAAAAGCAATTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1385
DB 1423 TCTAATTCAGATCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1482
QY 1386 GAGGAGAGGGGGGCGGCAAGAACTTTCACTTCAATTAAGGAGGAGGAGGAGGAGGAGG 1445
DB 1483 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1542
QY 1446 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1505
DB 1543 ATTGAAGAGATTAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1602
QY 1506 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1551

Db 1603 TCCACGACACACGTGTGTGAAAAACACGATACCCAAAGGAAACCAAT 1648

Search completed: December 5, 2003, 14:07:54
Job time : 463 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 13:51:42 ; Search time 3170 Seconds
(without alignments)
12351.581 Million cell updates/sec

Title: US-09-817-913-2
Perfect score: 1611
Sequence: 1 atgtcttg99gtctctgcgcgcg.....tccctcagttcttccccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pin:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_pmg:*
27: em_gse_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1152.6	71.5	1953	AK076080 Mus muscu
2	983.4	61.0	1201	BX425368 BX425368
3	943.2	58.5	1201	AL518877 AL518877
4	942.2	58.5	1201	BX421673 BX421673

5	935.4	58.1	1201	9	AL531607	AL531607
6	903	56.1	1201	9	AL546378	AL546378
7	901.4	56.0	1201	13	BX382026	BX382026
8	900	55.9	970	9	AL558916	AL558916
9	886.2	55.0	932	13	BQ925697	BQ925697
10	885.6	55.0	1201	13	BX333758	BX333758
11	884.2	54.9	951	13	BQ925328	BQ925328
12	872.2	54.1	1062	13	BQ276276	BQ276276
13	863.8	53.6	897	13	BQ642872	BQ642872
14	860.2	53.4	1201	9	AL541430	AL541430
15	857.4	53.2	1201	9	AL559849	AL559849
16	843.2	52.3	961	14	CD515346	CD515346
17	841.6	52.2	1060	12	BQ054741	BQ054741
18	841.2	52.2	992	13	BX458328	BX458328
19	836.2	51.9	919	13	BQ641927	BQ641927
20	833.6	51.7	906	9	AL546172	AL546172
21	831	51.6	910	9	AL522933	AL522933
22	820.4	50.9	1015	13	BQ064007	BQ064007
23	805.2	50.0	1072	12	BM923452	BM923452
24	797.4	49.5	1627	11	BC013141	BC013141
25	789.2	49.0	954	10	BG754715	BG754715
26	780.4	48.4	858	13	BU845191	BU845191
27	779.4	48.4	910	10	BG756668	BG756668
28	778	48.3	982	12	BM477303	BM477303
29	775	48.1	919	13	BQ423456	BQ423456
30	773.8	48.0	864	13	BQ521868	BQ521868
31	769.2	47.7	918	13	BQ679101	BQ679101
32	768.4	47.7	854	13	BU553962	BU553962
33	760.8	47.2	929	12	BQ057220	BQ057220
34	760.8	47.2	967	13	BQ680103	BQ680103
35	754	46.8	852	12	BQ218781	BQ218781
36	748	46.4	1025	12	BQ062801	BQ062801
37	745.6	46.3	872	13	BQ215409	BQ215409
38	737.4	45.8	778	12	BT727204	BT727204
39	736.8	45.7	1007	12	BM452740	BM452740
40	735.4	45.6	916	13	BU902967	BU902967
41	732.4	45.5	803	12	BG768807	BG768807
42	731	45.4	908	13	BQ885563	BQ885563
43	730.2	45.3	1031	12	BQ058662	BQ058662
44	729	45.3	1081	13	BQ069489	BQ069489
45	728.6	45.2	1034	12	BM472180	BM472180

ALIGNMENTS

RESULT 1
AK076080
LOCUS Mus musculus 11 days embryo whole body cDNA, RIKEN full-length
DEFINITION full insert sequence.
ACCESSION AK076080
VERSION AK076080.1 GI:26096626
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 9279253
MEDLINE 10349636
PUBMED 10349636
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
11042159

REFERENCE
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujimoto, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Ozaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
20530913
11076861

REFERENCE
AUTHORS
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotmann, M., Hume, D. A., Kamita, M., Lee, N. H., Lyons, P., Machionni, L., Mashima, J., Mazzarelli, D., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, K. P., Storch, K. F., Suzuki, H., Toyono, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, L., Wyshew-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohseki, S., and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

TITLE
JOURNAL
MEDLINE
PUBMED
21085660
11217851

REFERENCE
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
JOURNAL
REFERENCE
AUTHORS
6 (bases 1 to 1953)

ADACHI, J., AIZAWA, K., AKAHIRA, S., AKIMURA, T., AONO, H., ARAI, A., ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FUKUNISHI, Y., FURUNO, M., HANAGAKI, T., HARA, A., HAYASE, N., HIRAMOTO, K., HIRAO, T., HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., IZAWA, M., KASUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KOMNO, H., KOWADA, M., KOYA, S., KURIHARA, C., MATSUMURA, T., MIYAZAKI, A., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OKAZAKI, Y., OKIDO, T., OWA, C., SAITO, H., SAITO, R., SAKAI, C., SAKAI, K., SANO, H., SASEKI, D., SHIBATA, K., SHIBATA, Y., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., SHUKU, H., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TANAKA, T., TEJIMA, Y., TOYA, T., YAMAMURA, T., YAMANAKA, I., YASUNISHI, A., YOSHIDA, K., YOSHINO, M., MURAMATSU, M., and HAYASHIZAKI, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Shohri-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source
1. 1953
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2700094A20"
/db_xref="MGI:1910488"
/db_xref="taxon:10090"
/clone="2700094A20"
/issue_type="whole body"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="11 days embryo"
1. 1953
/note="histone deacetylase 1 (MGD|MGI:108086)"
/db_xref="MGI:108086"

misc_feature
BASE COUNT 531 a 468 c 466 g 468 t
ORIGIN

Query Match 71.5%; Score 1152.6; DB 11; Length 1953;
Best Local Similarity 88.0%; Pred. No. 6.8e-191;
Matches 1301; Conservative 0; Mismatches 174; Indels 4; Gaps 4;

126 GGACCCCGGAGAAAGCTCTTACTACTACGAGGGAGTTGAAATTACTTTATGA 185
18 GGAGCCAGAGAGAAAGCTCTTACTACTACGAGGGAGTTGAAATTACTTTATGA 77
186 CAAGGCCACCATGAAAGCTCAACCGAATCCGATGACTATTAATTGCTCACTAT 245
78 CAAGGCCACCATGAAAGCTCAACCGAATCCGATGACTATTAATTGCTCACTAT 137
246 GGTCTTACCGAAATGAAATCTATCGCCCTCACAAAGCCAAATGCTGAGAGATACC 305
138 GGCTCTACCGAAATGAGATCTACCGCTCTCACAAAGCCAAATGCTGAGAGATACC 197
306 AAGTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 365
198 AAGTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 257
366 GAGTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 425
258 GAATACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 317
426 CTGTTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 485
318 TTGTTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
486 AAGTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 545
378 AAGTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 437
546 GAGTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 605
438 GAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
606 TATCACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 665
498 TATCACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 557
666 GCTTCTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 725
558 GCTTCTACCAAGCGATGACTATTAATTGCTGCTCCATCCGTCGATTAACATGCTG 617
726 CAGGAATCGGGAGCTGATGAGGATTAACCGGGGCTGGCAAGCAAGTATTAAGCTTAA 785
618 CAGGAATCGGGAGCTGATGAGGATTAACCGGGGCTGGCAAGCAAGTATTAAGCTTAA 677
786 TACCCGCTCCGAGACGGATGATGAGGATTAACCGGGGCTGGCAAGCAAGTATTAAG 845
678 TACCCGCTCCGAGACGGATGATGAGGATTAACCGGGGCTGGCAAGCAAGTATTAAG 737
846 TCAGGAATGAGATGCTTCAAGCTGATGAGGATTAACCGGGGCTGGCAAGCAAGTATTA 905
738 TCAGGAATGAGATGCTTCAAGCTGATGAGGATTAACCGGGGCTGGCAAGCAAGTATTA 797

QY	906	CTATCTGGGGATCGGTTAGTTGCTTCATCTATCTATCAAGAGCAAGCCAAAGTGTG	965
Db	798	CTGTCTGGGAGCCGGTTAGTTGTTGCTTCATCTGAGCATC -AAGACAGCCAAAGTGTG	856
QY	966	GAATTGTTCAGAGACTTAACTTCGCTGATGCTGGAGGCGGTGTGTTACCAT	1025
Db	857	GAGTTTCGGAAGATTTCAACTTGCCCATGTGATGCTGGAGGAGGTGGTTACCATC	916
QY	1026	CGTAACGTTGCCGGGTGCTGACATATGACACAGCTGTGGCCCTTGATATCGAGATCCCT	1085
Db	917	CGGAATGGTGTCTGCTGCTGACATTACCAACAGCGGTGGCCCTTTCACACAGAGATCCCT	976
QY	1086	AATAGGCTTCATACCAATGACTTGAATCTTTGACCAAGATTTCAACTCCACATC	1145
Db	977	TATAG -TGCCCTACACAGACTTACTTGAATCTTTGACCGAGATATCTCTACATC	1035
QY	1146	AGTCCTTCATATATGACTTAACACAGAACAGATGAGTACTCTGAGAAATCAACAGCGA	1205
Db	1036	AGGCCCTTCACATATGACCAACAGAACCTAACAGATGACTCTGAGAAATATCAGGAGGCT	1095
QY	1206	CTGTTTGAAGACTTGAATCTGCTGCGACGACCTTGGGGTTCAAACGACGCGATTCT	1265
Db	1096	CTCTTTGAGAACTTGAGGATGCTGCCCATGCCCT -GGTTCAGATCAGAGCATCCCT	1154
QY	1266	GAGGACGCCATCCCTGAGAGAGTGGGATGAGACAGAAAGACCCCTGACAAAGCGATC	1325
Db	1155	GAGGATGCATCTCCCGAAGAGAGTGGGATGAGTGAAGTGAAGACCTTCGACAAAGCATC	1214
QY	1326	TGCATCTGCTCTCTGACAAACGAATTCCTGTGAGAAAGATTCTCCGATTTCTGAGAG	1385
Db	1215	TCCATCTGCTCTCTGACAAACGATTCCTGTGAGAAAGATTCTCCGATTCAGATGAG	1274
QY	1386	GAGGAGAGGGGGCCCGCAGAACTCTTCCAACTTCAAAAAAGCCAGAGAGTCAAAACA	1445
Db	1275	GAGGAGAGAGGTGCTCGCAAAACTCTTCTTAACCTTCAAAAAAGCCAAAGAGTTAAACA	1334
QY	1446	GAGGATGAAAAAGAAAGAACCCGACAGAGAGAAAGAAATATCCGAGAGAGAAAAACC	1505
Db	1335	GAGGATGAAAAAGAAAGAGATCTGMAAGAAAAAGAAAGTCAAGAAAGAGAAAAACC	1394
QY	1506	AAGAGAGAGAACCCAGAAAGCCAAAGGGGTCAAGAGAGAGGCCCAAGTTGGCTGAATGGAC	1565
Db	1395	AAGAGAGAGAACCCAGAAAGCCAAAGGGGTCAAGAAAGAGTCAAGTTGGCTTG -AGCAAG	1455
QY	1566	CTCTCAAGCTCTGCTTCCTGCTGAGTCCCTCAAGCTTTC	1604
Db	1454	GTCTGACAGCCCATCTTCTCCCAAGTTCTCACTTCTC	1492
RESULT 2	BX425368	1201 bp	linear EST 15-MAY-2003
LOCUS	BX425368		
DEFINITION	BX425368 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone		
VERSION	CL08B01.12H11.5-PRIME, mRNA sequence.		
ACCESSION	BX425368		
KEYWORDS	BX425368.1 GI:30770463		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., 'Jesse, J., and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 884.f For more information about this cluster, see http://www.genoscope.cns.fr/		

FEATURES					
Source					
	1. 1201				
	/organism="Homo sapiens"				
	/mol_type="mrna"				
	/db_xref="taxon:9606"				
	/clone="ClOB011ZH11"				
	/issue_type="NEUROBLASTOMA"				
	/clone_id="Homo sapiens NEUROBLASTOMA"				
	/note=Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector.				
	Library was not normalized."				
	Location/Qualifiers				
	1. 1201				
BASE COUNT	307 a	270 c	326 g	276 t	22 others
ORIGIN					
Query Match	61.0%;	Score 983.4;	DB 13;	Length 1201;	
Best Local Similarity	97.4%;	Pred. No. 2e-161;			
Matches 1018;	Conservative 11;	Mismatches 13;	Indels 3;	Gaps 3;	
Dy	97	GGAGGGGACCAAGATGGCCGACGCGAGGCCACC	CGAGGAAGTCGTCTACTACTACG	156	
Dy	90	GGAGGGGACCAAGATGGCCGACGCGAGGCCACC	CGAGGAAGTCGTCTACTACTACG	149	
Dy	157	ACGGGAGTGTGGAAATTACTATTATGACAAGGC	CACCAATAGACCTCACCGAATCC	216	
Dy	150	ACGGGAGTGTGGAAATTACTATTATGACAAGGC	CACCAATAGACCTCACCGAATCC	209	
Dy	217	GCATGACTCATTAATTTGCTGCTCACTATGCTCT	TCTACCGAAAATGGAATCTATGCCC	276	
Dy	210	GCATGACTCATTAATTTGCTGCTCACTATGCTCT	TCTACCGAAAATGGAATCTATGCCC	269	
Dy	277	CTCACAAGCCAATGCTGAGAAGATGACCAAATG	CCACAGAGTACTCACTTAAATCT	336	
Dy	270	CTCACAAGCCAATGCTGAGAAGATGACCAAATG	CCACAGAGTACTCACTTAAATCT	329	
Dy	337	TGCGCTCATCCGTCAGATTAACATGTCGGAGTA	CACACAAGCAGATCAGAGATTCAACG	396	
Dy	330	TGCGCTCATCCGTCAGATTAACATGTCGGAGTA	CACACAAGCAGATCAGAGATTCAACG	389	
Dy	397	TGCGTGAGGACTGTCAGATTTGCATGAGCCCTG	TTGAGTTCTGTCAGTTGTTACTGAGTG	456	
Dy	390	TGCGTGAGGACTGTCAGATTTGCATGAGCCCTG	TTGAGTTCTGTCAGTTGTTACTGAGTG	449	
Dy	457	GTTCTGTGGCAGTGTCTGTGAATACTTAATGAC	GCAGCAGCAGCAGATGCCGTGAATTTGG	516	
Dy	450	GTTCTGTGGCAGTGTCTGTGAATACTTAATGAC	GCAGCAGCAGCAGATGCCGTGAATTTGG	509	
Dy	517	CTGGGGGCGTGCACATGACAAAGAACGCCAGAT	CTGGCTTGTTAGTCGTAATGATA	576	
Dy	510	CTGGGGGCGTGCACATGACAAAGAACGCCAGAT	CTGGCTTGTTAGTCGTAATGATA	569	
Dy	577	TGCGCTTGAGCATCTTGAACCTGCTAAAGTATCA	CAGAGGCTGCTGTAATTGACATTG	636	
Dy	570	TGCGCTTGAGCATCTTGAACCTGCTAAAGTATCA	CAGAGGCTGCTGTAATTGACATTG	629	
Dy	637	ATATTTACCATGTGACCGCGTGAAGAAGCCCTT	CTACACCAAGACCGGCTCATGACTG	696	
Dy	630	ATATTTACCATGTGACCGCGTGAAGAAGCCCTT	CTACACCAAGACCGGCTCATGACTG	689	
Dy	697	TGTCCTTTCATTAATGAGAGATCTTCCAGGAA	CTGGGGAACCTACCGGATACCGGGG	756	
Dy	690	TGTCCTTTCATTAATGAGAGATCTTCCAGGAA	CTGGGGAACCTACCGGATACCGGGG	749	
Dy	757	CTGGCAAAAGCAGATATTAAGCTGTTAACTAAC	CGCTCCGAGACGGAGTTGATGACAGT	816	
Dy	750	CTGGCAAAAGCAGATATTAAGCTGTTAACTAAC	CGCTCCGAGACGGAGTTGATGACAGT	809	
Dy	817	CTATATGAGCCATTTTCAGCCGGCTCATGTC	CAAAAGTAAATGAGAGATGTTCCAGCCTTAGT	876	

Db 810 CCAATGAGGCGCATTTTCAAGCCGTCATGTCACAAAGTATGAGATGTTCCAGCCTAGTG 869
 Qy 877 CGGTGCTTCTTCAAGTGTGGCTCAGACTCCCTATCTGGGGATCGGTTAGTTCCTCAATC 936
 Db 870 CGGTGCTTCTTCAAGTGTGGCTCAGACTCCCTATCTGGGGATCGGTTAGTTCCTCAATC 929
 Qy 937 TATCTATCAAGAAGACAGCCCAAGTGTGTGAAATTTGTCAAGAGCTTTAACTGCTGCTATGC 996
 Db 930 TAACTATCAAGAAGACAGCCCAAGTGTGTGAAATTTGTCAAGAGCTTTAACTGCTGCTATGC 989
 Qy 997 TGATGTGGAGAGCGGTGTGTTACCACTTGTGAAAGTTGCGGTCGTGACATATATAGA 1056
 Db 990 TGATGTGGAGAGCGGTGTGTTACCACTTGTGAAAGTTGCGGTCGTGACATATATAGA 1048
 Qy 1057 CAGCTGTGGCCCTGATACGAGATCCCTATATGAGCTTCCATACAAATGACTACTTGAAT 1116
 Db 1049 CAGCTGTGGCCCTGATACGAGATCCCTATATGAGCTTCCATACAAATGACTACTTGAAT 1106
 Qy 1117 ACTTGGACGAGATTTCAAGCTCCA 1141
 Db 1107 ACTTGGACGAGATTTCAAGCTCCA 1131

RESULT 3
 AL518877 1201 bp mRNA linear EST 12-MAY-2003
 LOCUS AL518877 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
 DEFINITION CS0DA011YG11 5-PRIME, mRNA sequence.
 ACCESSION AL518877
 VERSION AL518877.2 GI:30537616
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 13, 2001 this sequence version replaced gi:12782370.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 884.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DA011AD060P1cluster=884.f. Contact :
 Feng Liang Email: fliang@life.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Parady Avenue Genoscope sequence ID : CS0DA011AD060P1.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DA011YG11"
 /tissue_type="NEUROBLASTOMA"
 /clone_lib="Homo sapiens NEUROBLASTOMA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was enriched
 with a NotI-oligo(dT) primer. Five prime end primed
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

BASE COUNT 315 a 263 c 319 g 282 t 22 others
 ORIGIN

Query Match 58.5%; Score 943.2; DB 9; Length 1201;
 Best Local Similarity 97.5%; Pred. No. 1.9e-154;
 Matches 967; Conservative 10; Mismatches 13; Indels 2; Gaps 2;

Qy 97 GAGAGCGAGCAAGATGGCGCAGACCGCAGGGCACCCGAGAGAAAGTCTTACTACTACG 156
 Db 74 GAGAGCGAGCAAGAGCGCAGACCGCAGGGCACCCGAGAGAAAGTCTTACTACTACG 133
 Qy 157 ACGGGATGTTGAAATTTACTATTATGAGCAAGGCCAACCCAAAGAGCCCTACCGAATCC 216
 Db 134 ACGGGATGTTGAAATTTACTATTATGAGCAAGGCCAACCCAAAGAGCCCTACCGAATCC 193
 Qy 217 GCATGACTCAATATTTGCTGCTCAACTATGCTCTTACCGAAAAATGAAATCTATGCC 276
 Db 194 GCATGACTCAATATTTGCTGCTCAACTATGCTCTTACCGAAAAATGAAATCTATGCC 253
 Qy 277 CTCACAAAGCCATGCTGAGAGATGACCAAGTACCAAGCATGACTACTAATTCT 336
 Db 254 CTCACAAAGCCATGCTGAGAGATGACCAAGTACCAAGCATGACTACTAATTCT 313
 Qy 337 TGGGCTCCATCCGTCAGATTAATATGTCGAGAGTACAGCAAGAGTGCAGAGATTCAAG 396
 Db 314 TGGGCTCCATCCGTCAGATTAATATGTCGAGAGTACAGCAAGAGTGCAGAGATTCAAG 373
 Qy 397 TTGGTGAAGACTGTCAGATTAATGATGAGCTGTTTGAATCTGTCAAGTTGTCTACTGTTG 456
 Db 374 TTGGTGAAGACTGTCAGATTAATGATGAGCTGTTTGAATCTGTCAAGTTGTCTACTGTTG 433
 Qy 457 GTTCTGTGCAAGTGTCTGTAATCTTAATATGAGACAGACGACATCCCGTGAATGG 516
 Db 434 GTTCTGTGCAAGTGTCTGTAATCTTAATATGAGACAGACGACATCCCGTGAATGG 493
 Qy 517 CTGGGGGCTGCAACATGCAAGAAGTCCGAGGCTTGTGCTTGTAGTCAATGATA 576
 Db 494 CTGGGGGCTGCAACATGCAAGAAGTCCGAGGCTTGTGCTTGTAGTCAATGATA 553
 Qy 577 TCGTCTTGGCCATCTGGAACCTGTAAGATGATACCAAGAGGTCGTCTACTTGAATTC 636
 Db 554 TCGTCTTGGCCATCTGGAACCTGTAAGATGATACCAAGAGGTCGTCTACTTGAATTC 613
 Qy 637 ATATTCACATGTGTGACGCGGTGAGAGAGGCTTTTACACACGAGCCGGTCAATGACTG 696
 Db 614 ATATTCACATGTGTGACGCGGTGAGAGAGGCTTTTACACACGAGCCGGTCAATGACTG 673
 Qy 697 TGTCTTTTCAATGATGAGAGTACTTCCAGAGAACTGGGGACCTTACGGGATACCGGGG 756
 Db 674 TGTCTTTTCAATGATGAGAGTACTTCCAGAGAACTGGGGACCTTACGGGATACCGGGG 732
 Qy 757 CTGGCAAGAAGCAATATTTATGCTTTACTACCGGCTCCGAGACGGGATGATGACGAGT 816
 Db 733 CTGGCAAGAAGCAATATTTATGCTTTACTACCGGCTCCGAGACGGGATGATGACGAGT 792
 Qy 817 CTAATGAGCCATTTTCAAGCCGGTCAATGTCACAAAGTATGAGATGTTCCAGCTATG 876
 Db 793 YCTATGAGCCATTTTCAAGCCGGTCAATGTCACAAAGTATGAGATGTTCCAGCTATG 852
 Qy 877 CGGTGCTTCAAGTGTGGCTCAGACTCCCTATCTGGGGATGCTGATGTTGTTCAAT 935
 Db 853 CGGTGCTTCAAGTGTGGCTCAGACTCCCTATCTGGGGATGCTGATGTTGTTCAAT 912
 Qy 936 CTAATCTCAAGAAGACAGCCCAAGTGTGGAATTTTCAAGAGCTTTAAGCTGCTATG 995
 Db 913 CTAATCTCAAGAAGACAGCCCAAGTGTGGAATTTTCAAGAGCTTTAAGCTGCTATG 972
 Qy 996 CTGATGTGGAGAGCGGTGTGTTACCACTTGTGAAAGTTTCCGCTGCTGACATATGAG 1055
 Db 973 CTGATGTGGAGAGCGGTGTGTTACCACTTGTGAAAGTTTCCGCTGCTGACATATGAG 1032
 Qy 1056 ACAAGTGTGGCCCTGATACGAGATCCCTAA 1087
 Db 1033 AAAGCTGTGGCCCTGATACGAGATCCCTAA 1064

RESULT 4
 BX421673 1201 bp mRNA linear EST 15-MAY-2003
 LOCUS BX421673 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
 DEFINITION BX421673 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA

Accession	Version	Keywords	Source	Organism	Reference Authors	Title	Journal	Comment			
BX421673	GI:30764107	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	Homo sapiens (human)	Homo sapiens	L.H.W.B., Gruber, C., Jeesee, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished	Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 884.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODH003BC05QDPI&cluster=884.f . Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODH003BC05QDPI.			
FEATURES	source	Location/Qualifiers	1..1201	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="CSODH003YE10"	/issue type="T CELLS (JURKAT CELL LINE)"	/cell_line="JURKAT CELL LINE"	/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"	/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT	ORIGIN	304 a	260 c	300 g	282 t	55 others					
Query Match	Best Local Similarity	58.5%; Score 942.2; DB 13; Length 1201;									
Matches 1008; Conservative	26; Mismatches 40; Indels 7; Gaps 5;										
Dy	97	GGAGGCGAGCAAAANTGGCCGACGACGCCACCAGGAAGAAGCTGTACTAATACTACG	156								
Dd	71	GAGGCGAGCAAAAGTAGCGCAGACGCAGCCACCAGGAAAAGTCGTACTAATACTACG	130								
Oy	157	ACGGGAGATGTTGAAAAATTACTATTATGAGACAAGGCCAACCAATGAAGCTCACCGAATCC	216								
Dd	131	ACGGGAGATGTTGAAAAATTACTATTATGAGACAAGGCCAACCAATGAAGCTCACCGAATCC	190								
Oy	217	GCATGACTCATTAATTTGCTGCTCAACATATGCTCTTACCGAAAAATGGAATCTATGCCC	276								
Dd	191	GCATGACTCATTAATTTGCTGCTCAACATATGCTCTTACCGAAAAATGGAATCTATGCCC	250								
Oy	277	CTCACAAGCCAAATGCTGAGAGATGACACATACACAGGATGATCACTAATTAATCT	336								
Dd	251	CTCACAAGCCAAATGCTGAGAGATGACACATACACAGGATGATCACTAATTAATCT	310								
Oy	337	TGCGCTCCATCCGTCCAGATTAACATGTGGAGATGACAGAACAGATGAGAGATTCACAG	396								
Dd	311	TGCGCTCCATCCGTCCAGATTAACATGTGGAGATGACAGAACAGATGAGAGATTCACAG	370								
Oy	397	TTGGTGAGAGCTGTCCAGATTTGAGATGAGCCGTGTTGAAGTTGTCAATTGCTGCTG	456								
Dd	371	TTGGTGAGAGCTGTCCAGATTTGAGATGAGCCGTGTTGAAGTTGTCAATTGCTGCTG	430								
Oy	457	GTTCTGAGAGAGCTGTGAAACCTTAATAGACAGACAGACATCGCCGCTGAATTTGGG	516								
Dd	431	GTTCTGAGAGAGCTGTGAAACCTTAATAGACAGACAGACATCGCCGCTGAATTTGGG	490								

Qy		517	CTGGGGGCGCTGCACCAATGCAAAGAAAGCCAGCATCTTGCCCTTGTGTAAGTCATAATATA	576
Db		491	CTGGGGGCGCTGCACCAATGCAAAGAAAGCCAGCATCTTGCCCTTGTGTAAGTCATAATATA	550
Qy		577	TGCTCTTGAGCCATCTCTGGAACCTGCTAAGAATACCAAGAGGCTGTGTACATTGACAATTG	636
Db		551	TGCTCTTGAGCCATCTCTGGAACCTGCTAAGAATACCAAGAGGCTGTGTACATTGACAATTG	610
Qy		637	AATATTCACCATGTTGACAGCGCTGGAAAAGGCCCTTTCTACACCACGACCGGGTCAATGACTG	686
Db		611	ATATTCACCATGTTGACAGCGGTGAAAAGAGCCCTTCTACACCACGACCGGGTCAATGACTG	670
Qy		697	TGTCCTTTCATAAGTATGAGAGAACCTTGCCACAGAACTGGGGAGCTACGGGATACCGGGG	756
Db		671	TGTCCTTTCATAAGTATGAGAGAACCTTGCCACAGAACT--GGGACCTACGGGATACCGGGG	729
Qy		757	CTGGCAAAAGACAAGTATTAATGCTGTTAACTTACCCGCTCCGAGACGGGATGATGACAGAT	816
Db		730	CTGGCAAAAGACAAGTATTAATGCTGTTAACTTACCCGCTCCGAGACGGGATGATGACAGAGT	789
Qy		817	CTATGAGGCCATTTTCAAGCCGCTCATGTGCCAAAGTAAATGAGATGTTCAAGCCTTAGTG	876
Db		790	CCTATGAGGCCATTTTCAAGCCGCTCATGTGCCAAAGTAAATGAGATGTTCCAGCCTTAGTG	849
Qy		877	CGGAGGCTTACAGATGAGGCTCACCTCCCTATCTGGGGATCGGTTAGTGTGCTTCATATC	936
Db		850	CGGAGGCTTACAGATGAGGCTCACCTCCCTATCTGGGGATCGGTTAGTGTGCTTCATATC	909
Qy		937	TATCTATCAAAGGACACGCCAAAGTGTG--GGAAATTTGTCAAGAGCTTTAACTCTGCCATG	995
Db		910	TAACTATCAAAGGACACGCCAAAGTGTGTTGGAAATTTGTCAAGAGCTTTAACTCTGCCATG	969
Qy		996	CTGATGCT--GGAGAGCGGTGTTTACACCACTTCGTAAAGTCCCGGTGCTGGAATATGA	1054
Db		970	CTGATGCTGGGAGAGCGGGGTTTACACCACTTCGTAAAGTTCSCGGGTGCTGGAATATGA	1029
Qy		1055	GACAGCGTGTGCCCCCTGGATACGAGATCCCTTAATGAGCTTCCATACATGACTACTTGTGA	1114
Db		1030	RACGAGCTGTGGCCCTTGGRAAMKRRPATICCMMAATARS--TTCCCTTCAATACACWCCTTT--	1085
Qy		1115	ATACTTTGAGACCAAGTATTCAAGTCCACATCAATGCTCCCTCAATATGACTCTAACAGAAAC	1174
Db		1086	AWMTTTTTTGCCCAATTTTCASCYCCMACAGTCTTCMAAWTATMAACHAAAMMRADARTCC	1145
Qy		1175	G	1175
Db		1146	K 1146	
RESULT 5				
AL531607				
LOCUS				
DEFINITION			1201 bp mRNA linear EST 23-MAY-2003	
VERSION			AL531607 Homo sapiens FETAL LIVER Homo sapiens CDNA clone	
KEYWORDS			CSO0M002THIS 5-PRIME, mRNA sequence.	
SOURCE			AL531607 GI:31069439	
ORGANISM			Homo sapiens (human)	
REFERENCE			Homo sapiens	
AUTHORS			Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
JOURNAL			1 (Bases 1 to 1201)	
COMMENT			L.M.B., Gruber,C., Jessee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished On Feb 13, 2001 this sequence version replaced gi:12795100. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 884.f For more information about this cluster, see	


```

OY 487 AGCAGACGACATCGCGTGAATTGGGCTGGGGCTTGACCAATGCAAGAACTCCG 546
|
|
|
Db 59 ATCAGAGACGACATCGCTGGAATTGGGCTGGGGCTTGACCAATGCAAGAACTCCG 118
|
|
|
OY 547 AGGCATCTGGCTTCTGTACGTCAATGATATGTCTTGGCATCTTGAACTGCTTAAT 606
|
|
|
Db 119 AGGCATCTGGCTTCTGTACGTCAATGATATGTCTTGGCATCTTGAACTGCTTAAT 178
|
|
|
OY 607 ATCACAGAGGGGCTGTACATTGATGATATTCACCATGGTGAAGGCTGGAAGAG 666
|
|
|
Db 179 ATCACAGAGGGGCTGTACATTGATGATATTCACCATGGTGAAGGCTGGAAGAG 238
|
|
|
OY 667 CCTTTCACACGACGACCGGCTCATGACTGTGCTCTTTCATAAGTATGAGAGTACTTCC 726
|
|
|
Db 239 CCTTTCACACGACGACCGGCTCATGACTGTGCTCTTTCATAAGTATGAGAGTACTTCC 298
|
|
|
OY 727 CAGGAACTGGGGACCTTACGGGATACGGGGCTGGCAAGACATATTAATGCTTTAACT 786
|
|
|
Db 299 CAGGAACTGGGGACCTTACGGGATACGGGGCTGGCAAGATATTAATGCTTTAACT 358
|
|
|
OY 787 ACCCGCTCGAGACGGGATGATGAGAGTCCATGATGAGGCAATTTCAAGCGGCTCATGT 846
|
|
|
Db 359 ACCCGCTCGAGACGGGATGATGAGAGTCCATGATGAGGCAATTTCAAGCGGCTCATGT 418
|
|
|
OY 847 CCAAGTATGAGAGATGTTCCAGCTAGTGGGCTGCTTACAGTGGCTGAGACTCC 906
|
|
|
Db 419 CCAAGTATGAGAGATGTTCCAGCTAGTGGGCTGCTTACAGTGGCTGAGACTCC 478
|
|
|
OY 907 TATCTGGGATGGGTTAGGTTGCTTCAATCTATCTATCAAGAGACGCAAGTGTGTG 966
|
|
|
Db 479 TATCTGGGATGGGTTAGGTTGCTTCAATCTATCTATCAAGAGACGCAAGTGTGTG 538
|
|
|
OY 967 AATTTGTCAAGAGCTTTAATCTGCGCTATGCTATGCTGGAAGGCGGTGTTACCATTC 1026
|
|
|
Db 539 AATTTGTCAAGAGCTTTAATCTGCGCTATGCTATGCTGGAAGGCGGTGTTACCATTC 598
|
|
|
OY 1027 GTTACCTTCCCGGCTCTGAGACATATGAGACAGCTGGCCCTGATGAGAGATCCCTA 1086
|
|
|
Db 599 GTTACCTTCCCGGCTCTGAGACATATGAGACAGCTGGCCCTGATGAGAGATCCCTA 658
|
|
|
OY 1087 ATGAGCTTCATACATGACTACTTGTAACTTTGGAACCAAGTTTCAAGCTTCACATCA 1146
|
|
|
Db 659 ATGAGCTTCATACATGACTACTTGTAACTTTGGAACCAAGTTTCAAGCTTCACATCA 718
|
|
|
OY 1147 GTTCTTCCAAATGACTTAACCGAAGACGAATGAGTACT -GAGAGAGTCAAAACGCGA 1205
|
|
|
Db 719 GTTCTTCCAAATGACTTAACCGAAGACGAATGAGTACTGAGGAGAAAGATCAAAACGCGA 778
|
|
|
OY 1206 CTGTTTGAGAACCTTTAGAAATGCTGCGCAAGCACT -GGGGTCCAAACGAGGCGATTC 1264
|
|
|
Db 779 CTGTTTGAGAACCTTTAGAAATGCTGCGCAAGCACTGAGGAGTCCAAATGAGGCGATTC 838
|
|
|
OY 1265 TGAAGACGCGATCCCTGAGAGAGTGGCGATGAGACGAAGCAAGCTTCGAAGCGCAT 1324
|
|
|
Db 839 TGAAGACGCGATCCCTGAGAGAGTGGCGATGAGACGAAGCAAGCTTCGAAGCGCAT 898
|
|
|
OY 1325 CTGATCTGCTCTCTGACAAAGAAATGCTGTGAGAAAGTTCCTCGATTCTGAAGA 1384
|
|
|
Db 899 CTGATCTGCTCTCTGACAAAGAAATGCTGTGAGAAAGTTCCTCGATTCTGAAGA 958
|
|
|
OY 1385 GAGAGGAGAGGGGGGCGGAGAACTCTTCCAACTTCAAAAAAGCAAGAGTCAAAAC 1444
|
|
|
Db 959 GAGAGGAGAGGGGGGCGGAGAACTCTTCCAACTTCAAAAAAGCAAGAGTCAAAAC 1016
|
|
|
OY 1445 AGAGATGAAAAAGAAAAAGACCCAGAGAG -GAGAAAAAGAAATCAACGAGAGAGAAA 1502
|
|
|
Db 1017 AGGTGAAAAATAGAAAGACAGAGAGAGAAAGAAATGATGATGAGAGAGAAA 1076
|
|
|
OY 1503 ACCAAGAGAGAGACCAAGAGCAAAAGGCTC 1535
|
|
|
Db 1077 RCCAGAGCAAAAGGGGSAAGAAAGATGAGGCGC 1109
|
|
|

```

```

RESULT 7
BX382026
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 884.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI073CF040P1ecluster=884.f. Contact :
Peng Liang Email : eliang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODI073CF040P1.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI073Y107"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="Just strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA
was digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT
299 a 262 c 325 g 282 t 33 others
ORIGIN
Query Match 56.0%; Score 901.4; DB 13; Length 1201;
Best Local Similarity 99.0%; Pred.No.3,6e-147;
Matches 924; Conservative 3; Mismatches 4; Indels 2; Gaps 2;

OY 116 GCAGACGACGAGCACCAGGAGGAAAGTCTGTACTACTACGAGGAGATGTTGAAATTA 175
|
|
|
Db 67 GCAGACGACGAGCACCAGGAGGAAAGTCTGTACTACTACTACGAGGAGATGTTGAAATTA 126
|
|
|
OY 176 CTATTATGAGCAAGGCCACCCATGAGAGCTTACCGAATCCGATGACTCATTAATTTGCT 235
|
|
|
Db 127 CTATTATGAGCAAGGCCACCCATGAGAGCTTACCGAATCCGATGACTCATTAATTTGCT 186
|
|
|
OY 236 GCTCAACTATGCTCTCTACCCGAAAAATGGAATCTATGSCCTCCCAAAAGCAATGCTGA 295
|
|
|
Db 187 GCTCAACTATGCTCTCTACCCGAAAAATGGAATCTATGSCCTCCCAAAAGCAATGCTGA 246
|
|
|
OY 296 GAGATGACCAAGTACCAAGCGATGATCACTTAATTTCTTGCCTCCATCCGTCAGAA 355
|
|
|
Db 247 GAGATGACCAAGTACCAAGCGATGATCACTTAATTTCTTGCCTCCATCCGTCAGAA 306
|
|
|
OY 356 TAAATGTCGAGTACAGCAAGCAAGTCAAGATTTCAAGCTTGGTGAAGACTGTCAGT 415
|
|
|
Db 307 TAAATGTCGAGTACAGCAAGCAAGTCAAGATTTCAAGCTTGGTGAAGACTGTCAGT 366
|
|
|
OY 416 ATTGATGGCTGTTTGAAGTCTGCAAGTGTCTAGTGTGCTGAGGCAAGGCTGT 475
|
|
|
Db 367 ATTGATGGCTGTTTGAAGTCTGCAAGTGTCTAGTGTGCTGAGGCAAGGCTGT 426
|
|
|
OY 476 GAACTTATATAGACAGACGACATCGCGTGAATTTGGGCTGGGGGCTGACCATC 535
|
|
|

```



```

Db 427 GAAATTAATAGCAGAGAGCAGCATCGCTGTGAATTTGGGCTGGGGCCCTGACCATGC 486
Qy 536 AAGAGATCCGAGGATCTGCTCTGTATGATGAATGATATGCTCTTGGCCATCCCTGGA 595
Db 487 AAGAGATCCGAGGATCTGCTCTGTATGATGAATGATATGCTCTTGGCCATCCCTGGA 546
Qy 596 ACTGCTAAAGTATCAGCAGAGGGTGTGTATGATGATGATGATGATGATGATGATGATG 655
Db 547 ACTGCTAAAGTATCAGCAGAGGGTGTGTATGATGATGATGATGATGATGATGATGATG 606
Qy 656 CTGGAAGAGGCTCTTCTACACACGAGACCGGGTATGATGATGATGATGATGATGATG 715
Db 607 CTGGAAGAGGCTCTTCTACACACGAGACCGGGTATGATGATGATGATGATGATGATG 666
Qy 716 AGATGATCTCCAGAGATGGGG-ACCTAGCGGATACCGGGCTGGGAAAGACAGATG 774
Db 667 AGATGATCTCCAGAGATGGGG-ACCTAGCGGATACCGGGCTGGGAAAGACAGATG 726
Qy 775 ATGCTGTTAACTACCCGCTCCGAGAGCGGATGATGATGATGATGATGATGATGATGATG 834
Db 727 ATGCTGTTAACTACCCGCTCCGAGAGCGGATGATGATGATGATGATGATGATGATGATG 786
Qy 835 AGCCGATCATGTCGAAAGTATGAGATGTTCCAGCTAGTGCCTGCTTACAGTGTG 894
Db 787 AGCCGATCATGTCGAAAGTATGAGATGTTCCAGCTAGTGCCTGCTTACAGTGTG 846
Qy 895 GCTCAGACTCCCTATCTGCGGATGCTGTTAGTGTCTTCAATCTATCTATCAAAGACAG 954
Db 847 GCTCAGACTCCCTATCTGCGGATGCTGTTAGTGTCTTCAATCTATCTATCAAAGACAG 906
Qy 955 CCAAGTGTGGAATTTGTCAAGAGCTTTAACTGCTGATGCTGATGCTGAGAGCGGTG 1014
Db 907 CCAAGTGTGGAATTTGTCAAGAGCTTTAACTGCTGATGCTGATGCTGAGAGCGGTG 966
Qy 1015 GTTACACCATTCGTAACGTTGCCCGGTGCTGGA 1047
Db 967 GTTACACCATTCGTAACGTTGCCCGGTGCTGGM 998

```

```

RESULT 8
AL558916 970 bp mRNA linear EST 31-MAY-2003
LOCUS AL558916 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ007YP20 5-PRIME, mRNA sequence.
ACCESSION AL558916
VERSION AL558916.2 GI:31283049
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 970)
L.I.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
unpublished
On Feb 15, 2001 this sequence version replaced gi:12903904.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
884.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ007DH100P1&cluster=884.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DJ007DH100P1.
Location/Qualifiers
1..970
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="CS0DJ007YP20"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 251 a 218 c 257 g 233 t 11 others
ORIGIN

```

```

Query Match 55.9%; Score 900; DB 9; Length 970;
Best Local Similarity 98.6%; Pred. No. 6.5e-147;
Matches 897; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 97 GGGAGCGAGCAAGATGCGCGCAGACCGAGGACCCGGAGGAAAGTGTGTTACTACTACG 156
Db 61 GGGAGCGAGCAAGATGCGCGCAGACCGAGGACCCGGAGGAAAGTGTGTTACTACTACG 120
Qy 157 AGCGGATGTTGGAATTTACTATTATGACCAAGGCCACCAATGAGCCACCGAATCC 216
Db 121 AGCGGATGTTGGAATTTACTATTATGACCAAGGCCACCAATGAGCCACCGAATCC 180
Qy 217 GCATGACTCAATATTGCTGCTCACTATGATGCTCTTACCGAAAAATGAAATCTATCGCC 276
Db 181 GCATGACTCAATATTGCTGCTCACTATGATGCTCTTACCGAAAAATGAAATCTATCGCC 240
Qy 277 CTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAACAGCATGATCAATTAATCT 336
Db 241 CTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAACAGCATGATCAATTAATCT 300
Qy 337 TGGGCTCCATCCCTCCAGATTAATGATGCTGAGATGCTGAGATGCTGAGATGCTGAG 396
Db 301 TGGGCTCCATCCCTCCAGATTAATGATGCTGAGATGCTGAGATGCTGAGATGCTGAG 360
Qy 397 TTGGTGAAGACTGCTCCAGATTTGATGAGCTGTTGATGATGATGATGATGATGATGATG 456
Db 361 TTGGTGAAGACTGCTCCAGATTTGATGAGCTGTTGATGATGATGATGATGATGATGATG 420
Qy 457 GTTCTGTGGCAAGTCTGTGAAAATTATTAAGCAGACAGACGATCCGCTGAATTTGG 516
Db 421 GTTCTGTGGCAAGTCTGTGAAAATTATTAAGCAGACAGACGATCCGCTGAATTTGG 480
Qy 517 CTGGGGGCTCGACCATGCAAGAGTCCGAGGACATCTGCTCTGTATGATGATGATGATG 576
Db 481 CTGGGGGCTCGACCATGCAAGAGTCCGAGGACATCTGCTCTGTATGATGATGATGATG 540
Qy 577 TCGCTTGGCCATCTCGAAGCTCTAAAGTATCAGCAGAGGCTGTGATGATGATGATG 636
Db 541 TCGCTTGGCCATCTCGAAGCTCTAAAGTATCAGCAGAGGCTGTGATGATGATGATGATG 600
Qy 637 ATATTCCATGATGAGACCGGCTGGAAGAGCCCTTCTTACACACGAGCCGGGTATGATG 696
Db 601 ATATTCCATGATGAGACCGGCTGGAAGAGCCCTTCTTACACACGAGCCGGGTATGATG 660
Qy 697 TGTCTTTTCAATGATGAGAGTACTTCCAGAGAACTGGGACCTTACCGGATTAACGGGG 756
Db 661 TGTCTTTTCAATGATGAGAGTACTTCCAGAGAACTGGGACCTTACCGGATTAACGGGG 720
Qy 757 CTGGCAAGCAAGTATTTATGTTTAACTACCGGCTCGAGACGGGATGATGATGATGATG 816
Db 721 CTGGCAAGCAAGTATTTATGTTTAACTACCGGCTCGAGACGGGATGATGATGATGATG 780
Qy 817 CTTATGAGGCAATTTTCAAGCCGCTGATGTCGAAAGTATGAGATGTTCCAGCTTATG 876
Db 781 CTTATGAGGCAATTTTCAAGCCGCTGATGTCGAAAGTATGAGATGTTCCAGCTTATG 840
Qy 877 CGGTGCTTACAGTGTGGTCAAGTCTCCATATCTGGGATTCGTTAGTGTGCTTCAATC 936
Db 841 CGGTGCTTACAGTGTGGTCAAGTCTCCATATCTGGGATTCGTTAGTGTGCTTCAATC 900

```


QY 937 TATCTATCAAGACACGCCAAGTGTGTGAATTTGTCAAGACCTTAACCTGCTATGC 996
 DB 901 TAACTATCAARAGACACSCCAAGTGTGTGAATTTGTCAAGACCTTAACCTGCTATGC 960
 QY 997 TGATGCTGGG 1006
 DB 961 TGATGCTGGG 970

RESULT 9
 B0926976 932 bp mRNA linear EST 20-AUG-2002
 LOCUS AGENCOURT_8804320 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6377915
 DEFINITION 5', mRNA sequence.
 ACCESSION B0926976
 VERSION B0926976.1 GI:22342007
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 932)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LHCW562 row: h column: 12
 High quality sequence stop: 668.
 Location/Qualifiers
 1. .932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6377915"
 /issue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_47"
 /note="Organ: Brain; Vector: pORF7, Site:1, XhoI, Site:2;
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 235 a 209 c 248 g 240 t
 ORIGIN

Query Match 55.0%; Score 886.2; DB 13; Length 932;
 Best Local Similarity 97.0%; Pred. No. 1.6e-144;
 Matches 903; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 286 CCAATGCTGAGAGATGACCAAGATGACAGAGATGACATTAATCTTGCGCTCCA 345
 DB 2 CAATGCTGAGAGATGACCAAGATGACAGAGATGACATTAATCTTGCGCTCCA 61
 QY 346 TCCGTCAGATTAACATGTCGAGATGACAGAGATGACAGATTAACATGTCGAGG 405
 DB 62 TCCGTCAGATTAACATGTCGAGATGACAGAGATGACAGATTAACATGTCGAGG 121
 QY 406 ACTGTCAGATTAACATGTCGAGATGACAGAGATGACAGATTAACATGTCGAGG 465
 DB 122 ACTGTCAGATTAACATGTCGAGATGACAGAGATGACAGATTAACATGTCGAGG 181

QY 466 CAAAGTCTGTGAACCTTAATAGCAGACGACATGCGCTGAATTTGGGCTGGGGCC 525
 DB 182 CAAAGTCTGTGAACCTTAATAGCAGACGACATGCGCTGAATTTGGGCTGGGGCC 241
 QY 526 TGCACCATGCAAGAAAGTCCGAGGCACTTGGCTTCTGTTACGTCAATGATTCGTTGG 585
 DB 242 TGCACCATGCAAGAAAGTCCGAGGCACTTGGCTTCTGTTACGTCAATGATTCGTTGG 301
 QY 586 CCATCTGGAACCTGCAAGATTAACACAGAGGGTGTGATGATTAATGATTCAC 645
 DB 302 CCATCTGGAACCTGCAAGATTAACACAGAGGGTGTGATGATTAATGATTCAC 361
 QY 646 ATGCTGACGCGGTGGAAGAGCGCTTCTACACAGACGAGCGGTCACTGTCCTTTC 705
 DB 362 ATGCTGACGCGGTGGAAGAGCGCTTCTACACAGACGAGCGGTCACTGTCCTTTC 421
 QY 706 ATGATGATGAGAGATCTTCCAGAGACGAGGAGCTTACGCGATACCGGGGCTGGCAAG 765
 DB 422 ATGATGATGAGAGATCTTCCAGAGACGAGGAGCTTACGCGATACCGGGGCTGGCAAG 481
 QY 766 ACAAGTATTAATGCTGTTAATACCGGCTCGAGAGCGGATGATGAGAGTCTTATGAG 825
 DB 482 GCAAGTATTAATGCTGTTAATACCGGCTCGAGAGCGGATGATGAGAGTCTTATGAG 541
 QY 826 CCATTTCAAGCGCGTCAATGTCGAAATGAGATGAGATGTCAGAGCTGAGTGGTCT 885
 DB 542 CCATTTCAAGCGCGTCAATGTCGAAATGAGATGAGATGTCAGAGCTGAGTGGTCT 601
 QY 886 TACAGTGTGCTCAGACTCCCTATCTGCGGATGCGTGAAGTGTCTTCAATCTATCA 945
 DB 602 TACAGTGTGCTCAGACTCCCTATCTGCGGATGCGTGAAGTGTCTTCAATCTATCA 661
 QY 946 AAGGACAGCGCAAGTGTGGAATTTGCAAGGCTTAACTGCTTAAGCTGAGTCTGG 1005
 DB 662 AAGGACAGCGCAAGTGTGGAATTTGCAAGGCTTAACTGCTTAAGCTGAGTCTGG 721
 QY 1006 GAGGCGGTGTTACACCATCTGTAAGTCCCGGCTGCTGACATATGAGACAGCTGG 1065
 DB 722 GAGGCGGTGTTACACCATCTGTAAGTCCCGGCTGCTGACATATGAGACAGCTGG 781
 QY 1066 CCCTGATACGAGATGCTCTTAATGAGCTTCCATACATGATCTTGAATCTTGGAC 1125
 DB 782 CCCTGATACGAGATGCTCTTAATGAGCTTCCATACATGATCTTGAATCTTGGAC 841
 QY 1126 CAGATTTCAAGCTCCACATCACTGCTTCCATATGATCTAACAGAACAGATGATACC 1185
 DB 842 CAGATTTCAAGCTCCACATCACTGCTTCCATATGATCTAACAGAACAGATGATACC 901
 QY 1186 TGGAGAAGATCAACAGAGAGCTGTTGAGAA 1216
 DB 902 CTGGAGAAGATCAACAGAGAGCTGTTGAGAA 932

RESULT 10
 BX333758 1201 bp mRNA linear EST 02-MAY-2003
 LOCUS BX333758
 DEFINITION CDNA clone CS0DD001Y16 5-PRIME, mRNA sequence.
 ACCESSION BX333758
 VERSION BX333758.1 GI:30333251
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segr@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 884.f. For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD001B808Q1&cluster=884.f>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DD001B808Q1.

FEATURES

SOURCE

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD001Y16"
/issue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 308 a 262 c 320 g 280 t 31 others
ORIGIN

Query Match 55.0%; Score 885.6; DB 13; Length 1201;
Best Local Similarity 97.0%; Pred. No. 2e-144;
Matches 939; Conservative 5; Mismatches 19; Indels 5; Gaps 4;

97 GGGAGGCGAAGAATGGCCGACGACGCGACCCGAGAAAGTCTGTACTACTACG 156
118 GGGAGGCGAAGAATGGCCGACGACGCGACCCGAGAAAGTCTGTACTACTACG 177
157 ACGGGAGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 216
178 ACGGGAGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 237
217 GCATGACTCATTAATTTGCTGCTCAATGCTCTCAACGAAATGAAATGAAATGCT 276
238 GCATGACTCATTAATTTGCTGCTCAATGCTCTCAACGAAATGAAATGAAATGCT 297
277 CTCACAAAGCCAAAGTCTGAGAGATGACCAAGATGACCAAGATGACCAAGATGAC 336
298 CTCACAAAGCCAAAGTCTGAGAGATGACCAAGATGACCAAGATGACCAAGATGAC 357
337 TGCCTCCATCCGTCAGATTAACATGTCGAGTACAGACCAATGCAAGATTCACG 396
358 TGCCTCCATCCGTCAGATTAACATGTCGAGTACAGACCAATGCAAGATTCACG 417
397 TTGGTGAAGACTGTCAGATTAATGATGATGATGATGATGATGATGATGATGATG 456
418 TTGGTGAAGACTGTCAGATTAATGATGATGATGATGATGATGATGATGATGATG 477
457 GTTCTGTGCGAAGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 516
478 GTTCTGTGCGAAGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 537
517 CTGGGGGCTCTGACCAATGCAAGATGCGAGGATCTGGCTTCTGTTACGTCATGATA 576
538 CTGGGGGCTCTGACCAATGCAAGATGCGAGGATCTGGCTTCTGTTACGTCATGATA 597
577 TCGTCTTGGCCATCTGGAACTGCTAAATGATCAACAGAGGATCTGTACATGACATG 636
598 TCGTCTTGGCCATCTGGAACTGCTAAATGATCAACAGAGGATCTGTACATGACATG 657
637 ATATTACCATGATGAGCGGCTGTAAGAGGCTTCTACACACGAGACCGGGTCACTG 696
658 ATATTACCATGATGAGCGGCTGTAAGAGGCTTCTACACACGAGACCGGGTCACTG 717
697 TGTCTTTCAATGAATGAGAGTACTTCCAGAGAACTGGGGAGCTACGGGATACCGGG 756
718 TGTCTTTCAATGAATGAGAGTACTTCCAGAGAACTGGGGAGCTACGGGATACCGGG 776
757 CTGGCAAGACAGATTAATGCTGTAACTACCGGCTCCGAGACGGGATGATGACGAGT 816
|||||

Db 777 CTGGCAAGCAAGTATTAATGCTGTAACTACACCGCTCCGAGACGGGATGATGACGAGT 836
Qy 817 CCTATGAGGCATTTTCAAGCCGATCATGTCCAAAGTATGAGATGTTCCAGCTAGTG 876
Db 837 CCTATGAGGCATTTTCAAGCCGATCATGTCCAAAGTATGAGATGTTCCAGCTAGTG 896
Qy 877 CGGTGTCTTACAGTGTGCTCAGACTCCCTATCTGGG-GATCGGTTAGTTCCTCAAT 935
Db 897 CGGTGTCTTACAGTGTGCTCAGACTCCCTATCTGGGCGATCGGTTAGTTCCTCAAT 956
Qy 936 CTATCTTCAAGAGACACCGCAAGTGTGAGATTTGTCAAGGC-TTTAACTGCTCAT 994
Db 957 CTATCTTCAAGAGACACCGCAAGTGTGAGATTTGTCAAGGC-TTTAACTGCTCAT 1016
Qy 995 GCGATGCTGGAGCGGCTGTACACATTCGTAAGTCCCGGCTGTCGACATATGA 1054
Db 1017 GCGATGCTGGAGCGGCTGTACACATTCGTAAGTCCCGGCTGTCGACATATGA 1074
Qy 1055 GACAGCTG 1062
Db 1075 MACGCTTG 1082

RESULT 11
BQ925328 951 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT 8837176 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6424758
DEFINITION 5', mRNA sequence.
ACCESSION BQ925328
VERSION BQ925328.1 GI:22340359
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 951)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gardar
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
<http://image.llnl.gov>
Plate: LCM2605 row: h column: 07
High quality sequence stop: 665.
Location/Qualifiers

FEATURES

SOURCE

1..951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6424758"
/issue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI, Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGACGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 236 a 216 c 253 g 238 t 8 others
ORIGIN

Query Match 54.9%; Score 884.2; DB 13; Length 951;
Best Local Similarity 96.5%; Pred. No. 3.7e-144;
Matches 920; Conservative 0; Mismatches 31; Indels 2; Gaps 2;

OY	298	AGATGACCAAGTACACAGCGCATGTACTCATTTAAATCTTGCGCTCCATCCGTCCAGTA	357
Db	1	AGATGACCAAGTACACAGCGCATGTACTCATTTAAATCTTGCGCTCCATCCGTCCAGTA	60
OY	358	ACATGTCCGAGTACAGCAAGCAGATTCAGAGATTCAACGTTGGTAGGAGCTGTCCAGTAT	417
Db	61	ACATGTCCGAGTACAGCAAGCAGATTCAGAGATTCAACGTTGGTAGGAGCTGTCCAGTAT	120
OY	418	TCGATGGCGCTTTGAGTCTGTCTGTCAAGTTGTCTACTGTGGTTCGTGTGCAAGTCTGTGA	477
Db	121	TCGATGGCGCTTTGAGTCTGTCTGTCAAGTTGTCTACTGTGGTTCGTGTGCAAGTCTGTGA	180
OY	478	AACCTTAATPAACAGACAGACGACATCCGCGTGAATTTGGGCGTGGGGGCGCTGCACCATGCAA	537
Db	181	AACCTTAATPAACAGACAGACGACATCCGCTGTGAATTTGGGCGTGGGGGCGCTGCACCATGCAA	240
OY	538	AGAAATCCGAGCACTTGGCTTCTGTTAACGTCAATGATATCGTCTTGGCCATCTCTGAAAC	597
Db	241	AGAAATCCGAGCACTTGGCTTCTGTTAACGTCAATGATATCGTCTTGGCCATCTCTGAAAC	300
OY	598	TGCTTAAAGTATCACACAGAGGCTGTACATTGACATTGATATTCACATGGTGAACGCG	657
Db	301	TGCTTAAAGTATCACACAGAGGCTGTACATTGACATTGATATTCACATGGTGAACGCG	360
OY	658	TGGAAGAGGCGCTTCTACACACACGAGCGGGTCAATGACGTGTCCTTTATPAAGTATGAG	717
Db	361	TGGAAGAGGCGCTTCTACACACACGAGCGGGTCAATGACGTGTCCTTTATPAAGTATGAG	420
OY	718	AGTACTTCCAGAACTGGGGAGCCTACCGGAGTACCGGGCTGGCAAGACAGATTAATG	777
Db	421	AGTACTTCCAGAACTGGGGAGCCTACCGGAGTATCGGGGCTGGCAAGACAGATTAATG	480
OY	778	CTGTTAACTAACCCGCTCCGAGACGGGATTTGATGACAGTCTTATGAGGCCATTTTCAAGC	837
Db	481	CTGTTAACTAACCCGCTCCGAGACGGGATTTGATGACAGTCTTATGAGGCCATTTTCAAGC	540
OY	838	CGGTCACTGTCCAAAGTATGAGAGATGTTCAAGCCTAGCGGGTGGTCTTAAGTGGCT	897
Db	541	CGGTCACTGTCCAAAGTATGAGAGATGTTCAAGCCTAGCGGGTGGTCTTAAGTGGCT	600
OY	898	CAGACTCCCTATCTGGGGATCGGTTAGGTGCTTCAATCTATCTATCAAGAGCACGCCA	957
Db	601	CAGACTCCCTATCTGGGGATCGGTTAGGTGCTTCAATCTATCTATCAAGAGCACGCCA	660
OY	958	AGTGTGTGAATTTGTCAAGACTTTAACTGTGCTATCTGATGCTGGAGGCGGTGTT	1017
Db	661	AGTGTGTGAATTTGTCAAGACTTTAACTGTGCTATCTGATGCTGGAGGCGGTGTT	720
OY	1018	ACACACTTCGTAAGGTTGCCGGGTCGGAGCATATGAGACAGCTGTGGCCCTGTATAGG	1077
Db	721	ACACACTTCGTAAGGTTGCCGGGTCGGAGCATATGAGACAGCTGTGGCCCTGTATAGG	780
OY	1078	AGATCCCTAATGAGCTTCCATACATGACTACTTGTGAATTACTTTGACACGATTTCAAGC	1137
Db	781	AGATCCCTAATGAGCTTCCATACATGACTACTTGTGAATTACTTTGACACGATTTTNCAGC	840
OY	1138	TCACATCAATGCTTTCCAAATATGACTTAAACAGAAACGAATGAGTACTGTGAGGAATCA	1197
Db	841	TCACATCAATGCTTTCCAAATATGACTTAAACAGAAATGAGTACTGTGAGGAATCA	899
OY	1198	AACAGCGCATGTTTGAAGAACTTGAATGTCGCGGACAGCACTGGGTTCCAA	1250
Db	900	AACAGCGCATGTTTGAAGAACTTGAATGTCGCGGACAGCACTGGGTTCCAA	951

RESULT 12	
BO276276	
LOCUS	
DEFINITION	BO276276 1062 bp mRNA linear EST 07-MAY-2002
ACCESSION	AGNCOCRT_6822351 NIH_MGC_111 Homo sapiens CDNA clone IMAGE:59319656
VERSION	5' mRNA sequence. BO276276.1 GI:20486484

SOURCE	KEYWORDS
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1062) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: CGAP (Stanford) CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LNCM214 row: c column: 06 High quality sequence stop: 682.
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
FEATURES	
source	1..1062 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5931965" /rissue_type="carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC lib1" /note="Organ: prostate; Vector: pOT7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dt printing; Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGACGAG(G). Library constructed by ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
BASE COUNT	276 a 248 c 284 g 254 t
ORIGIN	
Query Match	54.1%; Score 872.2; DB 13; Length 1062;
Best Local Similarity	95.9%; Pred. No. 4.4e-142;
Matches 939; Conservative	0; Mismatches 33; Indels 7; Gaps 4
Dy	97 GGGAGGGAGCAAGATGGCGCAGACGAGGAGCCGGAGGAAAGTCGTACTACTACG 156
Dd	13 GGGAGGGAGCAAGATGGCGCAGACGAGGAGCCGGAGGAAAGTCGTACTACTACG 72
Dy	157 ACGGGGATGTTGGAATAATTACTATTATGAGCACAGGCCAACCATGAAAGCTCACCGAATCC 216
Dd	73 ACGGGGATGTTGGAATAATTACTATTATGAGCACAGGCCAACCATGAAAGCTCACCGAATCC 132
Dy	217 GCATGATCTCATTAATTTTGCTGCTCAACTATGGTCTCTACCCGAAAATGGAATCTATCGCC 276
Dd	133 GCATGATCTCATTAATTTTGCTGCTCAACTATGGTCTCTACCCGAAAATGGAATCTATCGCC 192
Dy	277 CTCAACAAGCCAATGCTGAGAGATGACCAAAGTACCACAGGATGACTACATTAATCTT 336
Dd	193 CTCAACAAGCCAATGCTGAGAGATGACCAAAGTACCACAGGATGACTACATTAATCTT 252
Dy	337 TGCGCTTCATCCGTCACAGATPAACATGTCGAGAGTACAGCAAGCAAGCAGATGAGAATTCAAG 396
Dd	253 TGCGCTTCATCCGTCACAGATPAACATGTCGAGAGTACAGCAAGCAAGCAGATGAGAATTCAAG 312
Dy	397 TTGGTAGAGACTGTCCAGATATTCATGAGCGCTGTTGAGTTCTGTCAAGTTGTCTACTGGTG 456
Dd	313 TTGGTAGAGACTGTCCAGATATTCATGAGCGCTGTTGAGTTCTGTCAAGTTGTCTACTGGTG 372
Dy	457 GTTGTGAGGCAAGGCTGTGAAAACTTAATATAGCAGCAGACGAGCATCGCCGTGAATTGGG 516
Dd	373 GTTGTGAGGCAAGGCTGTGAAAACTTAATATAGCAGCAGACGAGCATCGCCGTGAATTGGG 432
Dy	517 CTGGGGGCTGCACCATGACAAAGAGTCGAGGCAATCGAGGATCTGCTGTAGCTCAATGATA 576

```

Db 433 CTGGGGGCTGCACATGCAAGAGTCCGAGGATCGCTTCTGTACGTATGATA 492
Qy 577 TCGCTTGGCCATCTCTGGAATCTGTAAGTATCACAGAGGCTGTACATTGACATG 636
Db 493 TCGCTTGGCCATCTCTGGAATCTGTAAGTATCACAGAGGCTGTACATTGACATG 552
Qy 637 ATATTACACATGATGACGCGCTGGAAGAGGCTTCTACACCAAGGAGCGGCTGATG 696
Db 553 ATATTACACATGATGACGCGCTGGAAGAGGCTTCTACACCAAGGAGCGGCTGATG 612
Qy 697 TGTCTTTTCAATAGTATGAGAGTACTTCCAGGAATCGGAGCTACGAGATACGGGG 756
Db 613 TGTCTTTTCAATAGTATGAGAGTACTTCCAGGAATCGGAGCTACGAGATACGGGG 672
Qy 757 CTGGCAAGACAAATATATGCTGTATCTACCTACCGCTCCGAGAGGATTTATACAGAGT 816
Db 673 CTGGCAAGACAAATATATGCTGTATCTACCTACCGCTCCGAGAGGATTTATACAGAGT 732
Qy 817 CCTATAGAGCCATTTTCAAGCGCGCTCATGTCCAAAGTATGAGATGTTCCAGGCTAGT 876
Db 733 CCTATAGAGCCATTTTCAAGCGCGCTCATGTCCAAAGTATGAGATGTTCCAGGCTAGT 792
Qy 877 CGG-TGGTCTTACAGTGTGCTCAGACTCCCTATCTGGGAGTCGGTTAGTTCCTCAT 935
Db 793 CGGTTGGTCTTACAGTGTGCTCAGACTCCCTATCTGGGAGTCGGTTAGTTCCTCAT 852
Qy 936 CTATCTATCAAAAGACAGCCCAAGTGTG--TGAAATTTGTCAAGACTTAACTGCTTA 993
Db 853 CTAACTATCAAAAGACAGCCCAAGTGTGAAATTTGTCAAGACTTAACTGCTTA 912
Qy 994 TGCTG-ATGCTGGAGGCGGTGTGTTACACATTCGTA--CGTTGCCGCTGTGACA 1049
Db 913 TGCTGATGCTGGAGGCGGTGTGTTAAACCAATCTTAACTGCTTCCCGGTGCTGAACA 972
Qy 1050 TATGAGACAGCTGTGCCC 1068
Db 973 TATGAGACAGCTGTGCCC 991

RESULT 13
B0642872 897 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT 8485556 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305494
DEFINITION B0642872.1 GI:21767044
ACCESSION B0642872.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 897)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Lou Stauch
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
http://image.llnl.gov
Plate: L1CM2527 row: n column: 23
High quality sequence stop: 654.
Location/Qualifiers
1..897
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6305494"

```

```

/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 223 a 207 c 245 g 221 t 1 others
ORIGIN
Query Match 53.6%; Score 863.8; DB 13; Length 897;
Best Local Similarity 98.4%; Pred. No. 1.3e-140;
Matches 882; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 421 ATGGCCCTGTTGAGTTCTGTCACTGTTCTACTGTTGTTCTGTGCAAGTGTGTAAC 480
Db 1 ATGGCCCTGTTGAGTTCTGTCACTGTTCTACTGTTGTTCTGTGCAAGTGTGTAAC 60
Qy 481 TTAATAAGCAGACGAGACATGCGGTGATTTGGCTGGGACCTGCACCATGCAAGA 540
Db 61 TTAATAAGCAGACGAGACATGCGGTGATTTGGCTGGGACCTGCACCATGCAAGA 120
Qy 541 AGTCCGAGATCTGCTTCTGTAGTCAATATATCGCTTGGCCATCTGGAAGTGC 600
Db 121 AGTCCGAGATCTGCTTCTGTAGTCAATATATCGCTTGGCCATCTGGAAGTGC 180
Qy 601 TAAAGTATCACAGAGGCTCTGTATGACATTTGATATTCACATGATGACGCGCTGG 660
Db 181 TAAAGTATCACAGAGGCTCTGTATGACATTTGATATTCACATGATGACGCGCTGG 240
Qy 661 AAGAGCCTTCTACACACGAGCCGGGTCACTGTGTCTTTCATAGTATGAGAGT 720
Db 241 AAGAGCCTTCTACACACGAGCCGGGTCACTGTGTCTTTCATAGTATGAGAGT 300
Qy 721 ACTTCCGAGAACTGGGAGCTTACCGGATACCGGGCTGGCAAGCAATATATGCTG 780
Db 301 ACTTCCGAGAACTGGGAGCTTACCGGATACCGGGCTGGCAAGCAATATATGCTG 360
Qy 781 TTAATACCGCTCCGAGAGCGGATTTGATGACAGTCTTATGAGCCATTTTCAAGCCGG 840
Db 361 TTAATACCGCTCCGAGAGCGGATTTGATGACAGTCTTATGAGCCATTTTCAAGCCGG 420
Qy 841 TCATGTCAAAGTATGAGATGTTCCAGCTTATGTCGTTTCACTGTCGCTCAG 900
Db 421 TCATGTCAAAGTATGAGATGTTCCAGCTTATGTCGTTTCACTGTCGCTCAG 480
Qy 901 ACTGCCATCTGGGAGATCGGTTAGTGTCTTCAATCTATCAAAAGACAGCCCAAGT 960
Db 481 ACTGCCATCTGGGAGATCGGTTAGTGTCTTCAATCTATCAAAAGACAGCCCAAGT 540
Qy 961 GTGTGAATTTGTCAAGAGCTTTAACTGCTTATGATGTCGAGAGGCGGTGTTACA 1020
Db 541 GTGTGAATTTGTCAAGAGCTTTAACTGCTTATGATGTCGAGAGGCGGTGTTACA 600
Qy 1021 CCATTGCTAAAGTTGCCCGGTGCTGACATATGAGACAGCTGTGGCCCTGATACGAGA 1080
Db 601 CCATTGCTAAAGTTGCCCGGTGCTGACATATGAGACAGCTGTGGCCCTGATACGAGA 660
Qy 1081 TCCTTATGAGCTTCCATTAAGATCTACTTTGAATCTTTGACCGAGATTTCAAGTCC 1140
Db 661 TCCTTATGAGCTTCCATTAAGATCTACTTTGAATCTTTGACCGAGATTTCAAGTCC 720
Qy 1141 ACATCAGTCTTCCAAATATGATTAACGAGAACGAAATGAGTACTGAGAAATCAAC 1200
Db 721 ACATCAGTCTTCCAAATATGATTAACGAGAACGAAATGAGTACTGAGAAATCAAC 780
Qy 1201 AGCAGCTGTTTGAAGACCTTAGATGTGCGGACGCACTGGGGTCAACGAGGCGA 1260

```

Db 781 AGGACCTGTTGAAACCTTAGATGCTGCCGACGACCACTGGGGGTCAAATGACAGCGA 840

QY 1261 TTCTTGAGGACGCGCATCCCTGAGAGAGTGG-CGATGAGGACGAGACGACCTCTGA 1315

Db 841 TTCTTGAGGACGCGCATCCCTGAGAGAGTGGCGGATGAGGACCAACGACCTCTGA 896

RESULT 14
AL541430 1201 bp mRNA linear EST 12-MAY-2003
LOCUS AL541430 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE006YL03
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL541430
VERSION AL541430.2 GI:30545601
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12872494.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Library: segrefgenoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 884.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE006CF020P1&cluster=884.f. Contact :
Peng Liang, Email: fliang@lifetech.com url :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE006CF020P1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE006YL03"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: PCWVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the PCWVSPORT 6 vector.
Library was not normalized."

BASE COUNT 294 a 250 c 312 g 284 t 61 others

ORIGIN

Query Match 53.4%; Score 860.2; DB 9; Length 1201;
Best Local Similarity 97.1%; Pred. No. 5.4e-140;
Matches 898; Conservative 8; Mismatches 16; Indels 3; Gaps 3;

QY 97 GGGAGCGGACGACGATGCGGCACAGCGGACACCGGAGAGAAAGTGTACTACTACG 156

Db 66 GGGAGCGGACGACGATGCGGCACAGCGGACACCGGAGAGAAAGTGTACTACTACG 125

QY 157 ACGGGATGTTGGAATTAATTAATTAATGACCAAGGACCAACCAATGAAGCTCAGCAATCC 216

Db 126 ACGGGATGTTGGAATTAATTAATTAATGACCAAGGACCAACCAATGAAGCTCAGCAATCC 185

QY 217 GCATGACTCAATATTTGCTGCTCACTAGTGTCTTACCGAAAAATGAAATCTATCGCC 276

Db 186 GCATGACTCAATATTTGCTGCTCACTAGTGTCTTACCGAAAAATGAAATCTATCGCC 245

QY 277 CTCACAAAGCGCAATGCTGGAGAGATGACCAAGTACCAAGCGATGACTCAATTAATTTCT 336

Db 246 CTCACAAAGCGCAATGCTGGAGAGATGACCAAGTACCAAGCGATGACTCAATTAATTTCT 305

QY 337 TGGCTTCATCCGTCAGATTAACATGTCGAGTACAGCAAGAGATGACAGATTCACAG 396

Db 306 TGGCTTCATCCGTCAGATTAACATGTCGAGTACAGCAAGAGATGACAGATTCACAG 365

QY 397 TTGGTGGAGACTGTCAGATTCAGATGAGCTGTTGAGTTCTGTGAGTTGTCTACTGTTG 456

Db 366 TTGGTGGAGACTGTCAGATTCAGATGAGCTGTTGAGTTCTGTGAGTTGTCTACTGTTG 425

QY 457 GTTCTGTGCAAGTGTGTAATACTTAATAGACAGACGACATCGCGGTAATTTGGG 516

Db 426 GTTCTGTGCAAGTGTGTAATACTTAATAGACAGACGACATCGCGGTAATTTGGG 485

QY 517 CTGGGGGCTGCAACCATGCAAGAAAGTCCGAGGCACTTGGCTTCTGTACGTATGANA 576

Db 486 CTGGGGGCTGCAACCATGCAAGAAAGTCCGAGGCACTTGGCTTCTGTACGTATGANA 545

QY 577 TGGTCTGGCCATCCGTAACCTGTAAGTATACACAGAGGGTGGCTTCAATTTGACTTG 636

Db 546 TGGTCTGGCCATCCGTAACCTGTAAGTATACACAGAGGGTGGCTTCAATTTGACTTG 605

QY 637 AATTCACCATGATGTCAGCGCTGGAAGAGCCCTTCTACACACGACCGGGTCACTAG 696

Db 606 AATTCACCATGATGTCAGCGCTGGAAGAGCCCTTCTACACACGACCGGGTCACTAG 665

QY 697 TGTCTTTCAATAGTATGAGAGTACTTCCAGAACTGGGG-ACCTACGGATACCGG 755

Db 666 TGTCTTTCAATAGTATGAGAGTACTTCCAGAACTGGGGACCTACGGATACCGG 725

QY 756 GTTGGGAAGACAAAGTATATCTGTATACTACCGGCTCCGACGAGATTATGACGAG 815

Db 726 GTTGGGAAGACAAAGTATATCTGTATACTACCGGCTCCGACGAGATTATGACGAG 785

QY 816 TCTATGAGGACCATTTTCAAGCGCGTCATGTCCTCAAGTATGAGATGTTCCAGCTAGT 875

Db 786 TCTATGAGGACCATTTTCAAGCGCGTCATGTCCTCAAGTATGAGATGTTCCAGCTAGT 844

QY 876 GCGGTGCTTCAAGTGTGCTCAAGCTCCTCATCTGAGGATCGTTAGTGTCTTCAAT 935

Db 845 GCGGTGCTTCAAGTGTGCTCAAGCTCCTCATCTGAGGATCGTTAGTGTCTTCAAT 904

QY 936 CTATCTATCAAGACACCGCAAGTGTGGAATTTGTCAAGCTTAACTGCTTATG 995

Db 905 CTATCTATCAAGACACCGCAAGTGTGGAATTTGTCAAGCTTAACTGCTTATG 963

QY 996 CTGATGCTGGAGGCGGCTTAC 1020

Db 964 CTGATGCTGGAGGCGGCTTAC 988

RESULT 15
AL559849 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL559849 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0D005Yg21 5-PRIME, mRNA sequence.
ACCESSION AL559849
VERSION AL559849.2 GI:31283980
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12905737.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Library: segrefgenoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 884.f For
more information about this cluster, see
http://www.genoscope.cns.fr/

cg1-bin/cluster.cgi?seq=CS0DG005AD110P1&cluster=884.f. Contact :
Feng Liang Email : fliang@life.technet.com URL :
http://fulllength.lifetech.com/ InVivoGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DG005AD110P1.
Location/Qualifiers

FEATURES

Source

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DG005Y621"

/issue_type="B CELLS (RAMOS CELL LINE)"

/cell_line="RAMOS CELL LINE"

/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(5') primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 323 a 263 c 322 g 263 t 30 others

ORIGIN

Query Match 53.2% Score 857.4; DB 9; Length 1201;

Best Local Similarity 97.5% Pred. No. 1.6e-139;

Matches 910; Conservative 3; Mismatches 14; Indels 6; Gaps 4;

```
OY 633 ATTGATATTCACATGTGAGCGCGTGGAGAGGCGCTCTACACCAAGCGGCGATG 692
DB 62 ATTGATATTCACATGTGAGCGCGTGGAGAGGCGCTCTACACCAAGCGGCGATG 121
OY 693 ACTGTGCTCTTCAATAGTATGAGAGTACTTCCAGGAAGTGGGAGCTTACGGGATAC 752
DB 122 ACTGTGCTCTTCAATAGTATGAGAGTACTTCCAGGAAGTGGGAGCTTACGGGATAC 180
OY 753 GGGGCTGGCAAGCAATATTAATGCTTAACTACCGGCTCCGAGACGGGATGATGAC 812
DB 181 GGGGCTGGCAAGCAATATTAATGCTTAACTACCGGCTCCGAGACGGGATGATGAC 240
OY 813 GAGTCTTATGAGGCGCATTTTCAAGCCGGTCAATGTCCTCAAGTATGAGATGTTCCAGCCT 872
DB 241 GAGTCTTATGAGGCGCATTTTCAAGCCGGTCAATGTCCTCAAGTATGAGATGTTCCAGCCT 300
OY 873 AGTGGGCTGCTTCAATAGTATGAGAGTACTTCCAGGAAGTGGGAGCTTACGGGATGAC 932
DB 301 AGTGGGCTGCTTCAATAGTATGAGAGTACTTCCAGGAAGTGGGAGCTTACGGGATGAC 360
OY 933 AATCTATATCAAGAGCAAGCCAGTGTGGAATTTTCAAGAGCTTTAACTGCTT 992
DB 361 AATCTATATCAAGAGCAAGCCAGTGTGGAATTTTCAAGAGCTTTAACTGCTT 420
OY 993 ATGCTGATGCTGGGAGCGGTGTTTACACATTCGTAAAGTTGCGCGGTGCTGACATAT 1052
DB 421 ATGCTGATGCTGGGAGCGGTGTTTACACATTCGTAAAGTTGCGCGGTGCTGACATAT 480
OY 1053 GAGACAGCTGTGGCCCTGTGATACGAGATCCCTAATGAGCTTCCATACATGACTACTT 1112
DB 481 GAGACAGCTGTGGCCCTGTGATACGAGATCCCTAATGAGCTTCCATACATGACTACTT 540
OY 1113 GAATATCTTGAACAGATTTCAAGTCCATCAATCACTCTTCAATATGACTTAACAGAAC 1172
DB 541 GAATATCTTGAACAGATTTCAAGTCCATCAATCACTCTTCAATATGACTTAACAGAAC 600
OY 1173 ACGAATGAGTACCTGGAGAGATCAACAGGAGCTTTGAGAACCTTAGAATGCTGCGG 1232
DB 601 ACGAATGAGTACCTGGAGAGATCAACAGGAGCTTTGAGAACCTTAGAATGCTGCGG 660
OY 1233 CACGCACTTGGGGTCCAAACGACGAGATTCCTGAGACGCGCATCCCTGAGAGAGTGGC 1292
DB 661 CACGCACTTGGGGTCCAAACGACGAGATTCCTGAGACGCGCATCCCTGAGAGAGTGGC 720
OY 1293 GATGAGAGCAAGAGACCTTGAAGAGGATCTGATCTGCTCTCTGACAAAGAAATT 1352
DB 721 GATGAGAGCAAGAGACCTTGAAGAGGATCTGATCTGCTCTCTGACAAAGAAATT 780
```

```
OY 1353 GCCTGTGAGAGAGTTCCTCCGATTCGAAAGAGAGAGAGGGGGCGCAAGAACTCT 1412
DB 781 GCCTGTGAGAGAGTTCCTCCGATTCGAAAGAGAGAGAGGGGGCGCAAGAACTCT 840
OY 1413 TCCAACTTCAAAAAGCCAAAGAGAGTCAAAAAGAGAGTAAAAAGAGAAAGCCAGAG 1472
DB 841 TCCAACTTCAAAAAGCCAAAGAGAGTCAAAAAGAGAGTAAAAAGAGAAAGCCAGAG 900
OY 1473 GAGAGAAAGAGATCAACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1532
DB 901 GAGAGAAAGAGATCAACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
OY 1533 GTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1565
DB 958 GTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
```

Search completed: December 5, 2003, 16:38:52
Job time : 3180 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 13:55:37 ; Search time 122 Seconds

(without alignments)
5828.428 Million cell updates/sec

Title: US-09-817-913-2

Perfect score: 1611

Sequence: 1 atgcctggggcctcgcgcgcg.....tccctcacgtcttcttcccc 1611

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTCTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1498	93.0	2111	1	US-08-528-255A-2
2	1498	93.0	2111	1	US-08-717-365-2
3	502.6	31.2	1943	3	US-09-282-305-7
4	502.6	31.2	1943	4	US-09-282-305-7
5	501.8	31.1	2019	3	US-09-282-305-5
6	501.8	31.1	2019	4	US-09-282-305-5
7	460.8	28.6	1576	3	US-09-282-305-9
8	460.8	28.6	1576	4	US-09-282-305-9
9	459	28.5	1826	3	US-09-282-305-1
10	459	28.5	1826	4	US-09-282-305-1
11	392.4	24.4	600	3	US-09-385-982-402
12	329.2	20.4	1475	3	US-09-282-305-3
13	329.2	20.4	1475	4	US-09-282-305-3
14	235.8	14.6	241	4	US-09-389-681-403
15	235.8	14.6	241	4	US-09-620-405B-403
16	235.8	14.6	241	4	US-09-433-826B-403
17	235.8	14.6	241	4	US-09-604-287A-403
18	204.6	12.7	841	3	US-08-998-416-317
19	122	7.6	242	4	US-09-016-434-612
20	112	7.0	283	4	US-09-113-294A-5130
21	80.8	5.0	290	4	US-09-113-294A-7484
22	77.8	4.8	212	4	US-09-213-294A-7066
23	77.2	4.8	7218	1	US-08-332-463-14
24	70.4	3.6	257	4	US-09-016-434-546
25	58.2	3.4	276	4	US-09-313-294A-5100
26	55.8	3.5	192	4	US-09-016-434-97
27	50	3.1	3489	2	US-08-728-323A-1

28	50	3.1	3489	4	US-09-298-568-1	Sequence 1, Appl1
29	50	3.1	3489	4	US-09-410-399-1	Sequence 1, Appl1
30	50	3.1	32207	2	US-08-770-379-20	Sequence 20, Appl1
31	50	3.1	32207	3	US-08-757-669A-20	Sequence 20, Appl1
32	50	3.1	32207	4	US-09-230-371A-20	Sequence 20, Appl1
33	49	3.0	1080	4	US-09-252-991A-1563	Sequence 1563, Ap
34	49	3.0	1779	4	US-09-252-991A-1553	Sequence 1553, Ap
35	49	3.0	1893	4	US-09-252-991A-1462	Sequence 1462, Ap
36	48.4	3.0	390	3	US-09-197-649-7	Sequence 7, Appl1
37	44.2	2.7	1926	4	US-09-249-585A-4	Sequence 4, Appl1
38	44.2	2.7	1931	2	US-09-130-114-2	Sequence 2, Appl1
39	43.8	2.7	3223	1	US-07-980-528-1	Sequence 1, Appl1
40	43.8	2.7	6755	3	US-08-931-999-4	Sequence 4, Appl1
41	42	2.6	480	2	US-08-726-306A-185	Sequence 185, App
42	42	2.6	90541	4	US-09-759-359A-3	Sequence 3, Appl1
43	41.8	2.6	5433	3	US-08-928-329-1	Sequence 1, Appl1
44	41.6	2.6	1042	3	US-09-276-531-77	Sequence 77, Appl1
45	41.6	2.6	43795	3	US-08-742-185-101	Sequence 101, App

ALIGNMENTS

RESULT 1
US-08-528-255A-2
; Sequence 2, Application US/08528255A
; Patent No. 5659016
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: FORUMAKA, YOICHI
; TITLE OF INVENTION: RPDL PROTEIN AND DNA
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: FLYNN, THIEL, BOUTELL & TANIS, P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/528,255A
; FILING DATE: September 14, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP6-227876
; FILING DATE: 22-SEPTEMBER-1994
; APPLICATION NUMBER: JP7-183763
; FILING DATE: 20-JULY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryance F. Chapman
; REGISTRATION NUMBER: 32549
; REFERENCE/DOCKET NUMBER: Futuya Case 1335
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2111
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal lung cDNA library
; FEATURE:

NAME/KEY: CDS
 LOCATION: 64..1512
 IDENTIFICATION METHOD: experimental examination
 US-08-528-255A-2

Query Match 93.0%; Score 1498; DB 1; Length 2111;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

97 GGGAGGGAGCAAGATGGCGGAGACGACGGGACCGGAGGAAGTCTGTTACTACTACG 156
 50 GGGAGGGAGCAAGATGGCGGAGACGACGGGACCGGAGGAAGTCTGTTACTACTACG 109
 157 ACCGGAGTGTGAATTTACTATTATGACACAGGCCACCCCAATGAAGCTTCACGAATCC 216
 110 ACCGGAGTGTGAATTTACTATTATGACACAGGCCACCCCAATGAAGCTTCACGAATCC 169
 217 GCATGACTCATTAATTTGCTGCTCAACTATGCTCTTACCGAAAAATGAAAATCTATGCC 276
 170 GCATGACTCATTAATTTGCTGCTCAACTATGCTCTTACCGAAAAATGAAAATCTATGCC 229
 277 CTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGGATGACTCACTTAATTTCT 336
 230 CTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGGATGACTCACTTAATTTCT 289
 337 TGCCTCCATCCCTCCAGATTAACATGTCGAGTACAGACAGACAGATGACAGATTTCAACG 396
 290 TGCCTCCATCCCTCCAGATTAACATGTCGAGTACAGACAGACAGATGACAGATTTCAACG 349
 397 TTGCTGAGAGACTGTCAGATTTTCATGAGCCCTGTTGAGTTCTGTCAGTTGCTTACTG 456
 350 TTGCTGAGAGACTGTCAGATTTTCATGAGCCCTGTTGAGTTCTGTCAGTTGCTTACTG 409
 457 GTTCTGAGAGAGTCTGTAAGTAACTTAATAGCAGACAGACAGATGCGCTGAAATGGG 516
 410 GTTCTGAGAGAGTCTGTAAGTAACTTAATAGCAGACAGACAGATGCGCTGAAATGGG 469
 517 CTGGGGGCTGACACATGACAAAGATCCGAGGATCTGCTTCTGTTACGTCATGATA 576
 470 CTGGGGGCTGACACATGACAAAGATCCGAGGATCTGCTTCTGTTACGTCATGATA 529
 577 TCGTCTTGGCCATCTGGAATCTGCTAAAGTATCAACAGAGGCTCTGTAACATGACATG 636
 530 TCGTCTTGGCCATCTGGAATCTGCTAAAGTATCAACAGAGGCTCTGTAACATGACATG 589
 637 ATATTACATGATGAGCGGCTGGAAGAGGCTTCTTACACACAGGACCGGCTCATGACTG 696
 590 ATATTACATGATGAGCGGCTGGAAGAGGCTTCTTACACACAGGACCGGCTCATGACTG 649
 697 TGTCTTTCTTAATGATGAGAGTACTCCAGAGAACTGGGGACCTTACGGGATACCGGGG 756
 650 TGTCTTTCTTAATGATGAGAGTACTCCAGAGAACTGGGGACCTTACGGGATACCGGGG 709
 757 CTGGCAAGAACAAGTATTATGCTGTTAACTTACCGGCTCCAGACGGGATTTGATGACAG 816
 710 CTGGCAAGAACAAGTATTATGCTGTTAACTTACCGGCTCCAGACGGGATTTGATGACAG 769
 817 CCTATGAGGCAATTTTCAAGCCGCTCATGTCAAAGTAAATGAGATGTTCCAGCTTACTG 876
 770 CCTATGAGGCAATTTTCAAGCCGCTCATGTCAAAGTAAATGAGATGTTCCAGCTTACTG 829
 877 CGGATGCTTCAAGTATGAGTCAAGTCCCTATCTGGGGAACTGGTTAGTCTTCAATC 936
 830 CGGATGCTTCAAGTATGAGTCAAGTCCCTATCTGGGGAACTGGTTAGTCTTCAATC 889
 937 TATCTATCAAGGACACGCAAGTGTGTAATTTGTCAGAGCTTTAACTGCTTATGC 996
 890 TAACTATCAAGGACACGCAAGTGTGTAATTTGTCAGAGCTTTAACTGCTTATGC 949
 997 TGAATCTGGAGGCGGCTGTTACCATTTCTTAAAGTTGCGGCTGCTGACATATGAGA 1056
 950 TGAATCTGGAGGCGGCTGTTACCATTTCTTAAAGTTGCGGCTGCTGACATATGAGA 1009

1057 CAGCTGTGCGCTGTGATACGAGATCCCTAATGAGCTTCCATCAATGACTACTTGAAT 1116
 1010 CAGCTGTGCGCTGTGATACGAGATCCCTAATGAGCTTCCATCAATGACTACTTGAAT 1069
 1117 ACTTTGAGCCAGATTTTCAAGCTTCAATCAATGCTTCTTCAATATGACTTACAGAACGA 1176
 1070 ACTTTGAGCCAGATTTTCAAGCTTCAATCAATGCTTCTTCAATATGACTTACAGAACGA 1129
 1177 ATGAGTACCTGGAGAAAGATGAAAACAGGACTGTTTGAACCTTAGAATCTGCGGACG 1236
 1130 ATGAGTACCTGGAGAAAGATGAAAACAGGACTGTTTGAACCTTAGAATCTGCGGACG 1189
 1237 CACTGGGGTCCAAACGACGAGCGATTTCTGAGACGCGCATCTGAGAGAGTGGCGATG 1296
 1190 CACTGGGGTCCAAACGACGAGCGATTTCTGAGACGCGCATCTGAGAGAGTGGCGATG 1249
 1297 AGACGAAAGCAACCTGACAAAGCGATTCGATTCGCTCTCTGACAAACGAATGGCT 1356
 1250 AGACGAAAGCAACCTGACAAAGCGATTCGATTCGCTCTCTGACAAACGAATGGCT 1309
 1357 GTGAGAGAGTCTCGATTTCTGAAGAGAGGAGAGGAGGCGGCAAACTCTTCCA 1416
 1310 GTGAGAGAGTCTCGATTTCTGAAGAGAGGAGAGGAGGCGGCAAACTCTTCCA 1369
 1417 ACTTCAAAAAAGCCAAAGAGTCAAAACAGAGATGAAAAAGAAAGACCCAGAGAGA 1476
 1370 ACTTCAAAAAAGCCAAAGAGTCAAAACAGAGATGAAAAAGAAAGACCCAGAGAGA 1429
 1477 AGAAGGATCACCAAGAGAGAGAAAAACAAGAGAGAACCAAGAGGAGGCTCA 1536
 1430 AGAAGGATCACCAAGAGAGAGAAAAACAAGAGAGAACCAAGAGGAGGCTCA 1489
 1537 AGAAGGAGCCAGTGGCGCTGATGAGACCTCTCCAGCTGCTGCTGAGTCCCT 1596
 1490 AGAAGGAGCCAGTGGCGCTGATGAGACCTCTCCAGCTGCTGCTGAGTCCCT 1549
 1597 CAGTTCTTTCCC 1610
 1550 CAGTTCTTTCCC 1563

RESULT 2
 US-08-717-365-2
 : Sequence 2, Application US/08717365
 : Patent No. 5763182
 :
 : GENERAL INFORMATION:
 : APPLICANT: NAKAMURA, YUSUKE
 : TITLE OF INVENTION: RPD, PROTEIN AND DNA
 : TITLE OF INVENTION: ENCODING THE SAME
 : NUMBER OF SEQUENCES: 2
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS, P.C.
 : STREET: 2026 Rambling Road
 : City: Kalamazoo
 : STATE: Michigan
 : COUNTRY: USA
 : ZIP: 49008-1699
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage
 : COMPUTER: IBM PC/XT/AT Compatible
 : OPERATING SYSTEM: MS-DOS 5.0
 : SOFTWARE: Wordperfect 5.0
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/717,365
 : FILING DATE: 23-SEP-1996
 : CLASSIFICATION: 536
 :
 : PRIORITY APPLICATION DATA:
 : APPLICATION NUMBER: 08/528 255
 : FILING DATE: September 14, 1995
 : APPLICATION NUMBER: JPE-227876
 : FILING DATE: 22-SEPTEMBER-1994
 : APPLICATION NUMBER: JPY-183763

FILING DATE: 20-JULY-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Terryence F. Chapman
 REGISTRATION NUMBER: 32549
 REFERENCE/DOCKET NUMBER: Furuya Case 1335
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (616) 381-1156
 TELEFAX: (616) 381-5465
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2111
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 IMMEDIATE SOURCE:
 LIBRARY: human fetal lung cDNA library
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 64..1512
 IDENTIFICATION METHOD: experimental examination
 US-08-717-365-2

Query Match 93.0%; Score 1498; DB 1; Length 2111;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

97 GGGAGCGAGCAAGATGGCGCAGACGAGGCGACCCGAGAGAAAGTCTGTACTACTAGC 156
 50 GGGAGCGAGCAAGATGGCGCAGACGAGGCGACCCGAGAGAAAGTCTGTACTACTAGC 109
 157 ACGGGAGATGTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 216
 110 ACGGGAGATGTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 169
 217 GCATGACTATATTTGCTGCTCACTATGCTCTCAACGAAATTAATTAATTC 276
 170 GCATGACTATATTTGCTGCTCACTATGCTCTCAACGAAATTAATTAATTC 229
 277 CTCACAAAGCCATGCTGAGAGATGACCAAGTACACAGCGATGACTAATTAATTC 336
 230 CTCACAAAGCCATGCTGAGAGATGACCAAGTACACAGCGATGACTAATTAATTC 289
 337 TCGGCTTCATCGCTCAAGATACATGTCGAGTACAGCAAGCATGACAGATTCACG 396
 290 TCGGCTTCATCGCTCAAGATACATGTCGAGTACAGCAAGCATGACAGATTCACG 349
 397 TTGGTAGAGACTGTCCAGATTTGATGAGGCTGTTGATGCTGCTCACTGCTG 456
 350 TTGGTAGAGACTGTCCAGATTTGATGAGGCTGTTGATGCTGCTCACTGCTG 409
 457 GTTCTGTGCAAGTGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 516
 410 GTTCTGTGCAAGTGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTC 469
 517 CTGGGGGCTGCAACATGCAAGAAGTCCGAGCATCTGCTTCTGTTACATGATTA 576
 470 CTGGGGGCTGCAACATGCAAGAAGTCCGAGCATCTGCTTCTGTTACATGATTA 529
 577 TCGTCTTGGCATCTCGGAACTGCTAAAGTATCAACAGAGGGTGTGATTAATGACATTC 636
 530 TCGTCTTGGCATCTCGGAACTGCTAAAGTATCAACAGAGGGTGTGATTAATGACATTC 589
 637 ATATTCACATGATGACGCGTGAAGAGGCTTCTTACACACAGGACCGGGTATGACTG 696
 590 ATATTCACATGATGACGCGTGAAGAGGCTTCTTACACACAGGACCGGGTATGACTG 649
 697 TGTCTTTCAATAGATGAGAGTACTTCCAGAGAACTGGGAACTTACGGGATACCGGGG 756
 650 TGTCTTTCAATAGATGAGAGTACTTCCAGAGAACTGGGAACTTACGGGATACCGGGG 709

757 CTGGCAAGACAGATTAATGCTGTTAACTACCCGCTCCGAGACGGGATTAATGACAGT 816
 710 CTGGCAAGACAGATTAATGCTGTTAACTACCCGCTCCGAGACGGGATTAATGACAGT 769
 817 CCTATGAGGCAATTTTCAAGCCGCTCATCTCCAAATTAATGAGATGTTCCAGCTTATG 876
 770 CCTATGAGGCAATTTTCAAGCCGCTCATCTCCAAATTAATGAGATGTTCCAGCTTATG 829
 877 CGGTGCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTAGGTGCTTCAATC 936
 830 CGGTGCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTAGGTGCTTCAATC 889
 937 TATCTATCAAGACACGCAAGTGTGAAATTTGTCAAGCTTTAACTGCTATG 996
 890 TAACTATCAAGACACGCAAGTGTGAAATTTGTCAAGCTTTAACTGCTATG 949
 997 TGATGCTGGAGGCGGTGCTTACCATTTCTTACGTTGCTCCGGTCTGACATTAAGA 1056
 950 TGATGCTGGAGGCGGTGCTTACCATTTCTTACGTTGCTCCGGTCTGACATTAAGA 1009
 1057 CAGCTGTGGCCCTGGATACGAGATCCCTAATGAGCTTCAATACATGACTTTGAAT 1116
 1010 CAGCTGTGGCCCTGGATACGAGATCCCTAATGAGCTTCAATACATGACTTTGAAT 1069
 1117 ACTTGGACCAATTTCAAGCTTCAACATGATCTTCAATTAATGACTAACAAGACGA 1176
 1070 ACTTGGACCAATTTCAAGCTTCAACATGATCTTCAATTAATGACTAACAAGACGA 1129
 1177 ATGAGTACTGGAAGAAATCAACAGCGACTGTTTGAACCTTGAATGCTGCGCAGC 1236
 1130 ATGAGTACTGGAAGAAATCAACAGCGACTGTTTGAACCTTGAATGCTGCGCAGC 1189
 1237 CAGCTGGGCTCAAAACGAGGCGATCTCTGAGAGCGCATCCCTGAGAGAGTGGGATG 1296
 1190 CAGCTGGGCTCAAAACGAGGCGATCTCTGAGAGCGCATCCCTGAGAGAGTGGGATG 1249
 1297 AGGACGAGACGACCTGCAAGCGATCTGATGCTCTCTGCAACGAATTCCT 1356
 1250 AGGACGAGACGACCTGCAAGCGATCTGATGCTCTCTGCAACGAATTCCT 1309
 1357 GTGAGGAGAGTCTTCGATTTGAAAGAGAGAGGAGGCGCAAGAACTCTTCCA 1416
 1310 GTGAGGAGAGTCTTCGATTTGAAAGAGAGAGGAGGCGCAAGAACTCTTCCA 1369
 1417 ACTTCAAAAAAGCCAAAGAGTCAAAAACAGAGATTAATAAGAAAGCCCAAGAGAGA 1476
 1370 ACTTCAAAAAAGCCAAAGAGTCAAAAACAGAGATTAATAAGAAAGCCCAAGAGAGA 1429
 1477 AGAAGAGATCAACCGAAGAGAGAGAAACCAAGAGAGAGCAAGAGCCAAAGGGGTCA 1536
 1430 AGAAGAGATCAACCGAAGAGAGAGAAACCAAGAGAGAGCAAGAGCCAAAGGGGTCA 1489
 1537 AGAGAGAGCCAAAGTGTGCTGAAATGACCTTCCAGCTCTGCTTCTGCTGATTCCT 1596
 1490 AGAGAGAGTCAAGTGTGCTGAAATGACCTTCCAGCTCTGCTTCTGCTGATTCCT 1549
 1597 CAGCTTCTTTTCCC 1610
 1550 CAGCTTCTTTCCC 1563

RESULT 3
 US-09-282-305-7
 ; Sequence 7, Application US/09282305
 ; Patent No. 6287843
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/282,305
 ; CURRENT FILING DATE: 1999-03-31

```

? PRIOR APPLICATION NUMBER: 607/080,563
? PRICE FILING DATE: 1999-04-03
? NUMBER OF SEQ. ID NOS.: 18
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 7
? LENGTH: 1943
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (57)..(1610)
JS-09-282-505-57

```

Query Match	31.2%	Score 502.6;	DB 3	Length 1943;
Best Local Similarity	65.5%	Pred No.3.5e-143;		
Matches 768;	Conservative 0;	Mismatches 399;	Indels 6;	Gaps 2

Oy	134	GAGGAAATCTGTTACTACTACGACGGGGATTTGTGAATTAATCTATTTAGACAGAGCCA	193
Db	122	GCGGGCGGTGTCCTACTTCTACAGACCGGATGTGGGCAACTACTACTACGAGGACAGGCA	181
Oy	194	CCCAATGAAGCCTCAACCGAATCCGATGACTATATAATTTGCTGTCAACTATGTCTCTA	253
Db	182	TCCGATGAAGCCGACCCCATCCGATAGCAGCACTCGCTGTGGCGGCTCAAGGCTCTCT	241
Oy	254	CCGAAATATGAAATCTATTCGGCTTCACAAAGCCAAATGCTAGAGAGATGACCAATACCA	313
Db	242	CAACCGATGACAGGGTACCGCCCAACCCCGGCCGACCGCGACTCTGTGGCTTCA	301
Oy	314	CAGCGATACTATCATTTAAATTTCTTGCGCTCCATCCGTCAGATTAATAGTCGAGATACAG	373
Db	302	CGCGACGACATCAATCAATCTTCTGTGGCTCCGTCAAGCCGAGAAAGCAGACGACAGAT	361
Oy	374	CAAGCAGATGCAGAGATTCAACGTTGTGTGAGGACTGTCAAGTATTCGATAGGCTGTTTGA	433
Db	362	CCGCTGTCTCAAGCCCTTCAACGTCGCGAGAGACTGCCCGCTTCGACGGGCTTACAG	421
Oy	434	GTTCTGTAGTTGCTACTCGGAGTTCTGTGGCAGATGCTGTGAACCTTAATTAACGACA	493
Db	422	CTTCTGCGAGACTATGCGGGGCGCTCCGTCCGGGGGCGCTCAAGCTTCAACACAGGCCA	481
Oy	494	GACGACATCGCCGTGAATTGGGCTGGGGCCCTGCACCATGCAAGAAAGTCCGAGGACATC	553
Db	482	---TGACATCGCAATCAATCGGTGCGGGGGCCTGCACACCGCAGAAAGTGGAGCGCTC	538
Oy	554	TGCGTTCTGTTAATGTCATATGATATGTCTTGGCCATCTCTGAACCTGCTAAAGATACCA	613
Db	539	GGGCTTCTGCACGTCGAATGACATGTGTCTGCCATATCTCGAGCTGCTCAACACATCACA	598
Oy	614	GAGGCTGCTGATACATTGACATTTGATATTCACCATAGTGTGACGGCGTGGAAAGAGGCTTCA	673
Db	599	GAGATTTCTGTATGTCGATATTCGATATTCACCATGTGTATGAGATGTGGAGAGGCTTTCTA	658
Oy	674	CACCAACGACCGGGTCATGACTGTGTCTTTCTATTAAGTATGAGAGATATCTTCCAGAAC	733
Db	659	CACAAACAGATAGGGTTATGACTGTCTCGTCCACAAGTTTGGTGAATTTTCCAGAAC	718
Oy	734	TGGGACTTACGGGATACCGGGGCTGGCAAAACAAAGTATTAAGTGTATTAATACCCGCT	793
Db	719	AGGGGATATCCGTGACATTTGGGCACTCAAAAGGAAGTACTCTCTGAATGTCCCTCT	778
Oy	794	CCGAGACCGGATTTGATGACGAGTCTTATAGGCCATTTTCAAGCCGCTCATGTCCAAAGT	853
Db	779	AGATGATGGATATGATGATGAAGTTACCAAGTCCCTTTTAAAGCAATATGGGAAAGT	838
Oy	854	AATGAGATGTTCCAGCCTAAGTGCAGGTGCTTCAAGATGTGGCTCAGACTCCATCTTGG	913
Db	839	TATGAGGTTTCCGCCCTGGTGCAAGTTGTCTCAGTGGTGTGATTCCTTGTCTGG	898
Oy	914	GGATCGGTAGGTTGCTTCATCTATCTATCAAAAGACACGCCAAGTGTGCGAATTTGT	973
Db	899	GGATAGGTTGGCTCTTCAACTTCAATCAAAAGTATAGGAGATGTGTAAGTATAT	958

QY	974	CAAGAGCTTTAACTGCGCTAATGCTGAATGCTGGAGAGGGATGGTTACACCAATTCGTAACT	1033
Db	959	GAGGCTTTCAACCTTCATCTTTGGCTCTTTGGTGGTGGATATATACATTAAGAAATGT	1018
QY	1034	TGCCCCGTCTGGAACATATGAGACAGCTGTGGCCCTTGATACGGAGATTCCTTAATGAGCT	1093
Db	1019	TGCACGCGCTGGTGGTTATGAGACGTGAGCTTCTTGGCCAMAGAGCCTTGAGACAAGAT	1078
QY	1094	TCCATACCAATGACTACTCTTGAATTAAGCTTTGACACGATTTTCAAGCTCCACATCAAGTCCCTTC	1153
Db	1079	GCGCTTATATGATGATCTATGAAATCTTCGGTCCAGATTACACTCTTATGTGGACCAAG	1138
QY	1154	CAATATGACTTAACAGAAACACGAATGAGTACCTGGAGAAAGATCAAACGAGACTGTTTGA	1213
Db	1139	TAACTATGAGAAACAAAATACACGACAAACATGGATGATATACACTTAAACTTCGTGA	1198
QY	1214	GAACCTTAGATGCTGGCCGACGACCTGGGGGTCCAACACGAGCGA---TTCCGTGAGA	1270
Db	1199	TAACTTTCAAAACTTGGACATGCTCCTAAGTGTCCACTTCAAGAGAGAGTTCCTGACAC	1258
QY	1271	CGCCATCCCTGAGAGAGATGGCCATATGAGAGA	1303
Db	1259	AGAAATACCTGAGCAAGATGAAGTACAAGATGA	1291

RESULT 4
US-09-883-720-7

```

/ Sequence 7, Application US/03863120
/ Patent No. 6479629
/ GENERAL INFORMATION:
/ APPLICANT: Baldwin, Donald A.
/ APPLICANT: Briggs, Steven P.
/ APPLICANT: Crane, Virginia C.
/ TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
/ FILE REFERENCE: 5718-44,
/ CURRENT APPLICATION NUMBER: US/09/883,720
/ CURRENT FILING DATE: 2001-06-18
/ PRIOR APPLICATION NUMBER: 09/282,305
/ PRIOR FILING DATE: 1999-03-31
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 7
/ LENGTH: 1943
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (57)..(1610)
US-09-883-720-7

Query Match          31.2%; Score 502.6; DB 4; Length 1943;
Best Local Similarity 65.5%; Pred. No. 3.5e-143;
Matches 768; Conservative 0; Mismatches 399; Indels 6; Gaps 2;

134 GAGGAAGTCTGTTACTACTACGACGGGGAGTGTGAAATTAATCATATATGAGACAAGGCCA 193
122 GCGGCGCGCTGTCTACTCTTCTACGACCCGGATGTGGCACTACTACTACACGGGCGGCCA 181
194 CCGAATGAGGCTCACCGGAATCGCGATGACTCATTAATTTGCTGCTCACTATGTCTCTTA 253
182 TCCATGTAAGCGCGACCGCATCCGGATGACGCACTCCGCTCTGGCGGCTACGGCCCTCT 241
254 CCGAAATATGSAATCTATGCGCCCTCTCAAAAGCCAATGCTGAGAGATGACCAAGTACCA 313
242 CAACCAAGTGAAGGTGTACGCGCCCAACCCGGCCCGCGACCGGCACTCTGCGGCTTCCA 301
314 CAGCGATGACTACATTAATTTCTTGCGCTCCATCCGTCAGATTAACATGTGCGAGTACAG 373
302 CGCGGAGACTACATCAACATTTCTGCGCTCGTCAAGCCGGAAACGACGAGGACAGAT 361
374 CAAAGCATGCGAGATTTCAACGTTGGTGAGACTGTCCAGTATTCGATGGCCCTGTTGA 433
362 CGCGCTGTCAAGGCTTCAACGTGCGCGAGGACTGCGCCGCTCTTCAAGCGCCCTTACAG 421

```

QY 434 GTTCTGTCAGTGTCTACTGAGTGTCTGTGCGAAGTGTGTGAATTAATTAAGCAGA 493
 DB 422 CTTCTCCGACGACATGCGGGGCGCTCCGCGGGGGGCGCTGAAGCTCAACGAGCCA 481
 QY 494 GACGACATCGCGCGTAATTGGGCTGGGCGCTGCAACCATGCAAGAAAGTCCGAGGATC 553
 DB 482 ---TGACATCGCAATCAACTGTGCGGGGGGCTGCAACGCGCAAGAAAGTCCGAGGATC 538
 QY 554 TGGCTTCTGTAGTCAATGATATGCTGCTGGCAATCTGGAACTGCTAAAGTATCAACA 613
 DB 539 GGGCTTCTGTAGTCAATGATATGCTGCTGGCAATCTGGAACTGCTAAAGTATCAACA 598
 QY 614 GAGGCTGCTGTAGTCAATGATATGCTGCTGGCAATCTGGAACTGCTAAAGTATCAACA 673
 DB 599 GAGAGTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
 QY 674 CACACGACGCGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
 DB 659 CACACGACGCGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
 QY 734 TGGGACCTACCGGATACCGGGGCTGGCAAGCAAGTATGATGATGATGATGATGATGAT 793
 DB 719 AGGGGATTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
 QY 794 CCGAGACGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853
 DB 779 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
 QY 854 AATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 913
 DB 839 TATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898
 QY 914 GATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 973
 DB 899 GATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
 QY 974 CAGAGCTTAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033
 DB 959 GAGGCTTTCACGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018
 QY 1034 TGGCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093
 DB 1019 TGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
 QY 1094 TCCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1153
 DB 1079 GCTGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
 QY 1154 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1213
 DB 1139 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
 QY 1214 GAACCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
 DB 1199 TATCTTCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1258
 QY 1271 CGCCATCCTGAGGAGATGAGGATGAGGAG 1303
 DB 1259 AGAATATCCTGAGCAAGATGAGATCAAGATGA 1291

RESULT 5
 US-09-282-305-5
 ; Sequence 5, Application US/09282305
 ; Patent No. 6287843
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/282,305

; CURRENT FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: 60/080,563
 ; PRIOR FILING DATE: 1998-04-03
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 2019
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (140) .. (1459)
 ; US-09-282-305-5
 Query Match 31.1%; Score 501.8; DB 3; Length 2019;
 Best Local Similarity 65.3%; Pred. No. 6.3e-143;
 Matches 753; Conservative 0; Mismatches 397; Indels 3; Gaps 1;
 QY 121 CCGAGGCAACCGGAGGAAGTCTGTACTACTACGAGGAGTGTGAAATTAATTAAT 180
 DB 189 CCGAGGCTGGAAGCGCGCGCTGTCTACTTCTACGACGGAGAGTGGCACTACT 248
 QY 181 ATGACAGGCAACCGAATGAAGCTCACCGAATCCGATGACTATTAATTGCTGCTA 240
 DB 249 ACGGGCAGGCGCAACCGATGAAGCGCACCGATCCGATGACCCAGCGCTGCGGC 308
 QY 241 ACTATGCTTCAACGAAATGAATATATGCGGCTTACAAAGCCATGCTGAGAGA 300
 DB 309 GCTACGGCTCTCTGACGAGATGAAGTGTCCGCTTACCTGCGCGACCGGACC 368
 QY 301 TGACCAATGACACAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 DB 369 TCTGCGCTTCCACGCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 428
 QY 361 TGTGAGATGACAGCAAGCATGACAGATGACAGATGACAGATGACAGATGACAGAT 420
 DB 429 AGCAGACCAAGATCCGCGGCTCAAGCGCTTCAAGCGCGAGGAGTCCCGCTTTCG 488
 QY 421 ATGCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 489 ACGGCTCTCAAGTTCGTGACAGATGACAGGAGGAGCTGTGTGCGCGCGCTGAC 548
 QY 481 TTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 DB 549 TCAACATGAGCA---TGATATGCGCATCACTGGGCGGCGGAGCTCACAGCCAA 605
 QY 541 AGTCGAGGATCTGGCTTCTGTATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 606 AGTGGAGGCTCCGGGTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 665
 QY 601 TAAATGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 DB 666 TCAAGTACCAACAGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 725
 QY 661 AAGAGGCTTTCACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 DB 726 AGAGGCTTTCATACACAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 785
 QY 721 ACTTCCAGAACTGGGAGCTTACGAGATACCGGAGTGGCAAGATTAATGATG 780
 DB 786 ATTCCCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 845
 QY 781 TTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 DB 846 TGAATGTTCCCTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 905
 QY 841 TCATGTCAAAGTAATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 906 TAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 965
 QY 901 ACTCCCTATCTGGGATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 966 ATTCAATGTCGGGTGACAGGTTGGGCTTTCAACCTCTCTAATTAAGGTCACAGGAT 1025

```

QY 961 GTGTGAATTGTCAAGAGCTTTAACTGCTATGATGCTGGAGCGGTGTACA 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1026 GTGTGAATTGTCAAGAGCTTTAACTGCTATGATGCTGGAGCGGTGTACA 1085
QY 1021 CCATTCGTAACGTTGCCGCTGTGACATATGAGACAGCTGTGCTGTGATACGAGA 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1086 CCATTAAGAAAGCTTGACAGGCTGTGTGCTACGAGACAGAGGAGTTGCCCTGTGATAGC 1145
QY 1081 TCCTTAATGAGCTTCCATACATGACTTCTTGAATCTTTGAGACGATTTCAAGTCC 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1146 TCATGTACAGATGCCACCTATAGTACTATGAGTATTTGGTCCAGATTTACACTGTAC 1205
QY 1141 ACATCAGCTCTTCCAAATATGACTTAACAGAAACAGCAATGAGTACCTGAGAAAGATCAAC 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1206 ATGTGCTCCCAAGTAACTGAGAAATTAATAACACAGGCACTCAATGAGATCAATAAAT 1265
QY 1201 AGCGACTGTTTGAAGACCTTAAGATGCTGCGACGCACTGGGGTCCAAACGCAAGCGA 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1266 CAATACTTCTGATTAATCTTTCAAACTCCGACATGCTCTAGTGTTCAGTTTCAGAGC 1325
QY 1261 TTCTTGAGAGCGC 1273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1326 GACCTCCTGAGGC 1338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6 US-09-883-720-5 Sequence 5, Application US/09883720

```

; Patent No. 6479629
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/883, 720
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/282,305
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (140)..(1459)
US-09-883-720-5

```

Query Match 31.1%; Score 501.8; DB 4; Length 2019;
Best Local Similarity 65.3%; Pred. No. 6.3e-143;
Matches 753; Conservative 0; Mismatches 397; Indels 3; Gaps 1;

```

QY 121 CGGAGGCAACCGGAGGAAAGTCTGTTACTACTACGAGCGGGAGTGGAAATTAATTAATT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 CGGAGGCTCAAGCGCGCGCTGCTACTTCTTACGACGCGGAGGTGGCACTACTACT 248
QY 181 ATGACACAGGCGCACCAATGAAAGCTCAACGAAATCCGATGACTCATTAATTTGTGCTCA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 ACGGGCAGGCGCACCGATGAAAGCGGACCGCATCCGATGACCCAGCGGTGCTCGGCC 308
QY 241 ACTATGCTCTTACCGAAATATGAAATCTATGCGCCTCAACAAAGCCAAATCTGAGAGA 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 GCTACGGCTCTTCAACAGATGAAAGTTCGCGCTTCAACCCCGCCGACCGGACCC 368
QY 301 TGACCAAGTACACAGGATGACTACTTAATTTCTTGGCTCCATCGCTCCAGATACACA 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 TCTGCGCTTCAAGCGGACGATTAAGTCTCTTCCGCTCCGCTCAACCCCGAAAGCC 428
QY 361 TGTGAGTACAGCAAGCAGATGACAGATTCAGAGTTCAAGTTGTGAGAGACTGTCCAGTATTCG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 429 AGCAGACAGATCCGCGGCTCAAGCGCTTCAACCTCGGCGAGAGCTCCCGCTCTTCG 488
QY 421 ATGAGCTGTTTGAGTCTGTGCACTGTCTACTGAGTGTCTGTGGCAAGTGTGTGAAC 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 ACGGTCTTACAGTCTTGTGCAAGCGTACCGGGGGGCTGTGTGTGGCGGCGGTCAAGC 548
QY 481 TTAATAAGCAGACGAGACATCCGCTGTAATTTGGCTGTGGGGCTGTGACATGCAAGA 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 549 TCAACCATGGCCA---TGATATGCCATCAACTGGGCGCGGAGCTTCCACACGCCAAGA 605
QY 541 AGTCCGAGCATGTGGCTCTGTGTAAGTCAATGATATTCGTTGGCATCTGGAACTGC 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 AGTGGAGGCTCCGGGTTTGTATGTTATGACATGTCTCGGCATCTTGAGCTCC 665
QY 601 TAAAGTATCACAGAGGCTGTGATACATGACATTAATTTACCAATGATGAGCGGTGG 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 666 TCAAGTACACAGCGCGCTTCTGTAGCTGACATTAATTCACACAGGGGAGGGGTGG 725
QY 661 AAGAGGCTTCTACACCAACGAGCCGGGTCACTGACTGTGCTTTCAATAGATGAGAGT 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 726 AGGAGGCTTTTATACACAGACCGGGGTGATGACAGTCTCATTTCCAAAGTTTGAGATT 785
QY 721 ACTTCCAGGAACTGGGGACCTTACGGGATACCGGGGCTGGCAAGCAAGATTAAGCTG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 786 ATTTCCCTGGGACAGGGGACATTCGTGATGTTGGGCACTCAAGGGTAAATTAATTAAT 845
QY 781 TTAATACCGGCTCCGAGACGGGATGATGACAGTCTTATGAGGCAATTTCAAGCGCG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 846 TGAATGTTCCCTGAGACAGATGATTAATGATGATGAGCTACAGTGTGTTCAAGCAA 905
QY 841 TCAATGCCAAAGTAAATGAGAGATTTCCAGCTAGTGGGTGTTCAAGTGTGGCTCAG 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 906 TAATGGGCAAGTGAATGAGAGTCTTCAACCTGTGCAAGTGTGCTCAAGTGTGGCGG 965
QY 901 ACTCCCTATGTGGGAGTGGTGTGCTTCAATCTATATCAAAAGCAAGCCAAAT 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 966 ATTCAATGTGGGTGACAGGTGTGGGTGTTCAACCTCTATTAAGGTGACAGAAAT 1025
QY 961 GTGTGAATTTGCAAGAGCTTTAACTGCTTACCTGCTGATGCTGGAGCGGTGTACA 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1026 GTGTGAATTTGAGAGTCTTCAAGCTCCGCTGTGCTGTGGTGTGGGTGTGATATA 1085
QY 1021 CCATTCGTAACGTTGCCGCTGTGACATATGAGACAGCTGTGGCTTGTGATACGAGA 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1086 CCATTAAGAAAGTTGACAGCGGTGTGGTGTACAGACAGAGATGTGCTGTGATAGC 1145
QY 1081 TCCTTAATGAGCTTCCATACATGACTACTTGAATTAATTTGAGACCAAGATTTCAAGTCC 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1146 TCATGTACAGATGCCACCTTAATGAGTACTATGATATTTGGTCCAGATTTACACTGTAC 1205
QY 1141 ACATCAGTCTTCCAAATATGACTTAACAGAAACAGATGAGTACTGAGAAATCAAAC 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1206 ATGTGCTCCCAAGTAACTGAGAAATTAATAACACAGGCACTCAATGAGATCAATAAAT 1265
QY 1201 AGCGACTGTTTGAACCTTGAATGATGTGCGGACGCACTGGGGTCCAAACGAGCGA 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1266 CAATACTTGAADAAATCTTTCAAACTCCGACATGCTCTAGTGTTCAGAGGC 1325
QY 1261 TTCTTGAGAGCGC 1273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1326 GACCTCCTGAGGC 1338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 7
US-09-282-305-9
Sequence 9, Application US/09282305
Patent No. 6287843
GENERAL INFORMATION:
APPLICANT: Baldwin, Donald A.
APPLICANT: Briggs, Steven P.
APPLICANT: Crane, Virginia C.
TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
FILE REFERENCE: 5718-44,

CURRENT APPLICATION NUMBER: US/09/282,305
 CURRENT FILING DATE: 1999-03-31
 PRIOR APPLICATION NUMBER: 60/080,563
 PRIOR FILING DATE: 1998-04-03
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: Patent Ver. 2.0
 SEQ ID NO 9
 LENGTH: 1576
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (38) .. (1336)
 US-09-282-305-9

Query Match 28.6%; Score 460.8; DB 3; Length 1576;
 Best Local Similarity 62.0%; Pred. No. 1,8e-130;
 Matches 764; Conservative 0; Mismatches 462; Indels 6; Gaps 2;

159 GGGGATGTTGGAATACTATTATGAGCAAGGCCACCATGAAAGCCTCAGCAATCCGC 218
 80 GGGATGTTGGCAATGTTCTTATGAGCAATCCCATGAGCAATGCTCTGT 139
 219 ATGACTCATATTTGCTGCTCAATATGCTCTACCGAAAAATGAAATCTATCCCT 278
 140 ATGACATCACTGTTGTTCTTATGATGACCTTCAATCAAAAGATGAGATATAGGCCA 199
 279 CACAAAGCCATGCTGAGAGATGACCAAGTACCAAGCATGATCAATTAATTTCTG 338
 200 CACAAAGCATATCCAAATAGAGCTTGCATATTCATTTGCTGATTAATGGAATTTCT 259
 339 CGCTCATCGTCCAGATTAATCATGTGAGTACAGCAAGCAAGATGAGATTCACAGTT 398
 260 CACCGATATCTCTGATTCACGACATATGACAGTACATTAATGATTAATCTT 319
 399 GGTGAGACTGTCAGATTTGATGAGCTGTTGAGTTCTGAGTTGCTACTGCTGT 458
 320 GGGAGAGACTGTCGGCTTTGATTAATTTGTTGAGTTGCGCAATCTATGCGGAGGA 379
 459 TCTGTGCAAGTCTGTGAATTAATAGCAGACAGAGCATCGCGTGAATGGCT 518
 380 ACTTTAGATGCTGCTGCGAGATTAATCATTAATATGTCATTTGCTGATTAATGGCT 439
 519 GGGGCTGAGCATGCAAAAGATGCGAGCATGCTGCTGATACGATGATATC 578
 440 GGTGAGTATATATGCAAAAGATGAGGCTTCAAGCTTCTGTTACATTAATGATCTA 499
 579 GTCTTGCCATCTGGAATGCTTAAGTATCACAGAGGCTGCTGATCAATGATGAT 638
 500 GTATTGGAATTTGAGGCTTCTCAAGTACCAAGGCTTCTTATTAATGATGAT 559
 639 ATTACCATGATGAGCGGCTGGAAGAGGCTTCTACACAGCAAGCGGCTGATGATG 698
 560 GTCCATCATGAGATGAGATGAAGAGGCTTTATTTATTTCACTGACAGGATATGATG 619
 699 TCTTTTATATGATGAGA---GTACTTCCAGAGAACTGGGAGCTACAGGATACCGG 755
 620 AGTTTCCACAAGATGATGAGCTGTTCTTCTGGAACAGGTATTAAGATATAGGA 679
 756 GCTGCAAAAGACAGTATTAAGCTGTTAACTACCGGCTCGAGACGGGATTTGATGAG 815
 680 GAAAGGAGGAAATTTATGCTATACATTCCTTAAAGATGAGATGATGACAT 739
 816 TCTATAGGCAATTTTCAAGCGGCTGATGCAAGATTAAGATGATGATGATGAT 875
 740 AGCTTACTGCGCTTTTAAACAATATGCAAGATTTGAGACATATGCTGCTGT 799
 876 GCGGTGCTTAAGTGTGCTGAGATCCCTATCTGGGATGAGTATGATGATGAT 935
 800 GCTATGTTCTTAAGTGTGCTGAGATCCCTATCTGGGATGAGTATGATGATGAT 859
 936 CTATCTATCAAGAGACAGCGCAAGTGTGGAATTTGTCAAGAGCTTAACTGCTATG 995

860 CTCTCTATGAAGGCCATGCTGAATGTGTAAAGTTTGTCAAGAAATTCATATCCCTT 919
 996 CTGATGCTGGAGGGGCTGTATACCAATTCGTAACGTTGCGGCTGCTGACATATGAG 1055
 920 CTGTATCACTGAGAGTGTGTGATACCAAGAGAAATGATGACAGGCTGTGTGAA 979
 1056 ACAGCTGTGGCCCTGATACGAGATTCCTTAATGAGCTTCCATACATGATCTTTGAA 1115
 980 ACTGGGCTCTTTTATGACACAGACATCCCAATATGATTCAAAAAATATATATTTAG 1039
 1116 TACTTTGACCAATTTTCAAGTCCATCAATCTCTTCAATATGATCAACAGCAAG 1175
 1040 TACTTTGCTCCAGATTTATCATTTGAAAGTTCCAAATTTGACATGACATTTTGAAC 1099
 1176 AATGATGATCTGAGAAATCAACAGCACTGTTTGAACCTTGAATGCTGCGCAC 1235
 1100 AAGACTATCTCAATGATCAATCAAGTGAAGATGAGAGATTTGGGATCATGACAT 1159
 1236 GCACCTGGGCTCAACAGCGAGATTCCTGAGAGCGCATCCCTGAGAGAGTGGCAT 1295
 1160 GCTCCGTGTTCAATGCAAGAGGTTCTCCGATTTTATATCCCGACTTT---GAT 1216
 1296 GAGAGCAAGAGACCTGTAAGAGCATGATGCTCTCTGTAAGAAAGATTTGCC 1355
 1217 GAGATGAATGAGATCTGATGAGAGCTGTTGACCAAGACATTCAGACAGAGATTCAC 1276
 1356 TGTGAGAGAGATTCCTCGATTCGTAAGAGA 1387
 1277 CGTATGATGATGATGATGAGAGTGA 1308

RESULT 8
 US-09-883-720-9
 ; Sequence 9, Application US/09883720
 ; Patent No. 6479629
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/883,720
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 09/282,305
 ; PRIOR FILING DATE: 1998-03-31
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 1576
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (38) .. (1336)
 ; US-09-883-720-9

Query Match 28.6%; Score 460.8; DB 4; Length 1576;
 Best Local Similarity 62.0%; Pred. No. 1,8e-130;
 Matches 764; Conservative 0; Mismatches 462; Indels 6; Gaps 2;

159 GGGGATGTTGGAATACTATTATGAGCAAGGCCACCATGAAAGCCTCAGCAATCCGC 218
 80 GGGATGTTGGCAATGTTCTTATGAGCAATCCCATGAGCAATGCTCTGT 139
 219 ATGACTCATATTTGCTGCTCAATATGCTCTACCGAAAAATGAAATCTATCCCT 278
 140 ATGACATCACTGTTGTTCTTATGATGACCTTCAATCAAAAGATGAGATATAGGCCA 199
 279 CACAAAGCCATGCTGAGAGATGACCAAGTACCAAGCATGATCAATTAATTTCTG 338
 200 CACAAAGCATATCCAAATAGAGCTTGCATATTCATTTGCTGATTAATGGAATTTCT 259
 339 CGCTCATCGTCCAGATTAATCATGTGAGTACAGCAAGCAAGATGAGATTCACAGTT 398


```

Db      | 260 CACCGGTAATCTCTGATTCACGACCACTATATGCAATGCACTAATCACTAATCAATCTT 319
Qy      | 399 GGTGAGGACCTGTCAGATTCGATGAGCGCTGTTGAGTCTGCTGCTGCTGCTGCTGCTGCT 458
Db      | 320 GGAGAAAGCTGTCGGCTTGTGATTAATTTGTTGATTTGTTGATTTGTTGATTTGTTGATTTG 379
Qy      | 459 TCTGTGCAAGTGTGTGAACTTAATATAGACAGACGCAATCGCGGTGAATTTGGCT 518
Db      | 380 ACTTATATGCTGCTGCAAGTTAAATCAATTAATATGATGATGATGATGATGATGATGATGATG 439
Qy      | 519 GGGGGCTGCAACATGCAAGAAAGTCCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
Db      | 440 GGTGGCTCAATCAATGCAAAAGTGTGAGGCTTCAGGCTTCTGTTACATTAATGATCTA 499
Qy      | 579 GTCTGGCAATCTCGGAATGCTTAAGTATACAGAGGGGTGCTGATCAATGATGAT 638
Db      | 500 GTATTTAGAAATCTGAGGCTTCTCAAGTACCATCCAGGGTCTTTATATTTAGACATTTGAT 559
Qy      | 639 ATTCAACATGCTGACGCGGTGGAAGAGGCTTCTTACACACAGACCGGCTCATGCTGTG 698
Db      | 560 GTCTATATGAGATGAGATTTGAAGAAAGCTTTATTTTCACTGACAGGATATGATCTGTG 619
Qy      | 699 TCCCTTCAATAGATGAGA---GTACTTCCAGAACTGAGGACCTTACGAGATACCGGG 755
Db      | 620 AGTTTCAACAGTATGAGTGAACCTGTTCTTCTGGAACAGGATGATTTAAAGATATAGGA 679
Qy      | 756 GCTGGCAAGACAGATTAATGCTGTTAATCAACCGCTCCGACAGCGGATGATGAGAGAG 815
Db      | 680 GAAAGGAAAGAAATATATGCTATCAACATTCACATTAAGATGGAATGAGATGAGACT 739
Qy      | 816 TCTATGAGGCACTTTTCAAGCCGCTCATGTCAAAGTAAATGAGATGTTTCAAGCTAGT 875
Db      | 740 AGCTTACTCGGCTTTTAAACAATATTTGCCAAAGTTGTTGAGACATATCTGCCCTGT 799
Qy      | 876 GCGGTGCTTTACAGTGTGCTCAAGCTCCCTATCTGAGGATCGGTTAGTGTCTTCAAT 935
Db      | 800 GCTATGTTCTTCAATGAGGCTGATGATTCATTTGCGAGGATCGTTTAAAGCTGCTTCAAT 859
Qy      | 936 CTATCTTCAAGAGACACGCAAGTGTGGAATTTTTCAGAGCTTTAATCTGCCATG 995
Db      | 860 CTCTCTATGAAAGCCATGCTGAATGTATAAGTTTTCAGAAATTCATATATCTCCCTT 919
Qy      | 996 CTGATGCTGAGAGCGGTGTTACACCATTCGTAAAGTTCGCGGTCTGACATATGAG 1055
Db      | 920 CTGTTACTGAGGTGTGTGATACACCAAGAGAAATGAGACAGGTGTGTGAGCTGTTGAA 979
Qy      | 1056 ACAGCTGTGCGCTGATACGAGATCCCTAATAGCTTCAATACATGACTTACTTTGAA 1115
Db      | 980 ACTGGGGTCTTTTGAACAGAACTCCCAATAGATTTCCAAAAATGAATATATTTGAG 1039
Qy      | 1116 TACTTTGACCAAGATTTTCAAGCTCCATCTCTTCCATATATGATCAATACCAAAACAG 1175
Db      | 1040 TACTTTGCTCCAGATTTATATTAATTAAGTTTCAATTTGAACATGACAAATTTTAAACGT 1099
Qy      | 1176 AATGAGTACTGGAAGATCAAAACAGGACCTGTTTGAAGACTTAAAGATCTGCCGAC 1235
Db      | 1100 AAGACTTACTCAATGATCAATCAAAAGTGAAGTGAAGATTTGCCGTATACATACAGAT 1159
Qy      | 1236 GCACCTGGGGTTCAAACGACGAGATTCCTGAGGACCGCATCCCTGAGAGAGTGGCGAT 1295
Db      | 1160 GCTCTGGTGTTCAAATGCAAGAGGTTCTCCGATTTTATATCTCCGAGCTTT---GAT 1216
Qy      | 1296 GAGACCAAGACGACCTTGAACAGGCACTTCGATCTGCTCTCTGACAAACGAATGCC 1355
Db      | 1217 GAAAGTGAATTTGATCTCGATGAACGTTGTAACAGCACACTCAAGAACAGCAATTCAC 1276
Qy      | 1356 TGTGAGGAAGATCTCTCGATTTCTGAAGAGA 1387
Db      | 1277 CTTGATGATGATCTATGAAGTGAACATGA 1308

```

RESULT 9

```

US-09-282-305-1
; Sequence 1, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 1:
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (29)..(1405)
US-09-282-305-1
Query Match 28.5%; Score 459; DB 3; Length 1826;
Best Local Similarity 61.8%; Pred. No. 7e-130;
Matches 729; Conservative 0; Mismatches 450; Indels 0; Gaps 0;
Qy      | 131 CCGAGAAAGTCTGTACTTACTTACGACGGGATGTTGAAATTAATTAATGAGCAAG 190
Db      | 88 CCGCCCGCGGTAGCTATTCTTACGACCGGTGATGAGACTACTACTACGCGCAAG 147
Qy      | 191 CCAACCAATGAACTTCACCGCAATCCGATGACTCAATTTGCTGCTCAATAGTCT 250
Db      | 148 TCAACCGATGAACCCACCGCATCCGAATGGGCACTGCTGTGTGCTCACTACGCGCT 207
Qy      | 251 CTAACGAAATGAAATCTATGCTCCCTCAAAAGCCATCTGAGAGATGACCAATGA 310
Db      | 208 CCAACGCTCTGAGGCTCCCGCCCTACCGGCTCTGAGGCGGACATCCGCGCTT 267
Qy      | 311 CCAACGAGATGATTAATTAATTTTGGCTCCATCCGTCGATTAATCAATGTGAGTA 370
Db      | 268 CCACTCGAGACTAGCTGCTTCTCGCTCGGCAACGAAACCGGCTGCTTGA 327
Qy      | 371 CACGACGATGACAGATTCACAGTGTGAGACTGTCAATATTCATGAGCTGTT 430
Db      | 328 CCGCGGCAATTAAGGCTTTAAGTGTGAGACTGCTGCTGCTGCTGCTGCTGCTGCT 387
Qy      | 431 TGAATCTGTCAATGCTGATCTGCTGTGTGCTGTGCAATGCTGTAATTAAGCA 490
Db      | 388 CCGCTTCTGCAAGGCTCCGCTGAGGGAACATCGCGCGCTCAAGCTTAAACGCGG 447
Qy      | 491 GCAGACGACATGCGCTGAATTTGGGTGGGGCTGCAACATGCAAGAAAGTCCAGGC 550
Db      | 448 GAGACGCAATCAACCTCACTGAGGCGGCGCTTCAACAGCCCAAGAAAGCGAGGC 507
Qy      | 551 ATGCGCTTCTGTTACGTCATGATATCTTGGGCACTCTGGAATCTGTAATGATCA 610
Db      | 508 CTCGAGTCTGCTAGCTCAACGATGATGATGATGATGATGATGATGATGATGATGAT 567
Qy      | 611 CCAAGGATCTGATCAATTAATGATTAATCAATGATGATGATGATGATGATGATGATG 670
Db      | 568 CAGCGGTGCTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627
Qy      | 671 CTACACACGAGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 730
Db      | 628 CTTCATCAAAACGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 687
Qy      | 731 AACTGGGACCTACCGGATACCGGGCTGCAAGCAATATATGATGATGATGATGATGATG 790
Db      | 688 TACTGACATATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 747
Qy      | 791 GCTCCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850

```



```

Db      748 CCGAGTGTATGTATGATGACACCACTTTCGTGCTGTTTCAATGCAATCATTTAAGAA 807
Qy      851 AGTAATGAGATGATGTCAGCTTAGTGCGGAGTCTTAAAGTGTGGCTCAGACCTCCATC 910
Db      808 AGTTATGAGATGATTAACAGCAACCTGATGTTCTTCCATGCGAGCTGACTCTTGGC 867
Qy      911 TGGGAGTCGTTAGTTAGTTCTTCAATCTATCTATCAAAAGACACGCCAAGTGTGTGAAT 970
Db      868 TGGAGACAGGTGATGTTGTTCAACCTGCTGTGAAAGGGTCATGTCGACCTCGGTTT 927
Qy      971 TGTCAAGAGCTTAACTGCTCTATGCTGATGCTGGAGCGGTGTGTTACCACTTGTTA 1030
Db      928 CCTTAGGTCTCAAAAGTTCTTATGATGTTTAAAGTGTGAGGTTTACCACTCAGAA 987
Qy      1031 CGTTGCCGGTGTGACATATAGACAGCTGTGGCCCTGAGTACCGAAGATCCCTATATA 1090
Db      988 TGTTCGACGCTGTGCTGCTACGAGACCGAGTGTGCTTGGAGTTGAACCTGATACAA 1047
Qy      1091 GCTTCCATACATGACTACTTGTGAATATCTTTGGACCAATTTCAAGCTCCACATCAGTCC 1150
Db      1048 GCTGCTTACAAATGATTAATGATGATGATCTTTGGCCCTGATTAATCTTCAATCCAAC 1107
Qy      1151 TTCCAAATGACTAACCAACACGAAATGATGATCTGGAGAGATCAAAACAGGACTGTT 1210
Db      1108 AAAAGTGTGAAAACCTGATATCCAAAGGACTTGGAGAACATTAAGAACATGATATT 1167
Qy      1211 TGAAGACTTGAATGCTGCGGACGACCTGCGGTCCAAACGAGGCAATTCCTAGAA 1270
Db      1168 GAGAACTGTGTCAAAAGATGAAATGATGTTCCAGACCTCAATTCATGACAGACCTGAGA 1227
Qy      1271 CGCCATCCCTGAGAGAGTGGCGATGAGAGCAAGACA 1309
Db      1228 CCTGAAGCTCCAGAGAGAAAGAGAGACATGACAA 1266

RESULT 10
US-09-883-720-1
: Sequence 1, Application US/09883720
: Patent No. 6479629
: GENERAL INFORMATION:
: APPLICANT: Baldwin, Donald A.
: APPLICANT: Briggs, Steven P.
: APPLICANT: Crane, Virginia C.
: TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
: FILE REFERENCE: 5718-44,
: CURRENT APPLICATION NUMBER: US/09/883,720
: PRIOR FILING DATE: 2001-06-18
: PRIOR APPLICATION NUMBER: 09/282,305
: PRIOR FILING DATE: 1999-03-31
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1826
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (29)..(1405)
US-09-883-720-1

Query Match      28.5%; Score 459; DB 4; Length 1826;
Best Local Similarity 61.8%; Pred. No. 7e-130;
Matches 729; Conservative 0; Mismatches 450; Indels 0; Gaps 0;

```

```

Qy      251 CTACGGAAAAATGAAATCTATCGCCCTCACAAAGCCAAATGCTGAGAGATGACCAAGTA 310
Db      208 CCAAGCGCTCTCTGAGACTCTCCGCCCTAACCCGGGCTCTGAGGCGGACATCCGCGCTT 267
Qy      311 CCAACCGATGACTATCAATTAATCTTGGGCTCCATCCCTCCAGATTAACATGTGCGAATA 370
Db      268 CCACTCCGACGACTAGTGGCTTCTCGGCTCCGACCCGAAACCCGGGTGCTCTGA 327
Qy      371 CAGCAAGCAGATGACAGATTCACAGTTGATGAGAGCTGTCCAGATTTGATGAGGCTGTT 430
Db      328 CCGCGCGCCATTAAGCCGTTTAACGTGGTGAAGACATGCCCCGTGTGACGCTCTT 387
Qy      431 TGAATCTGTAGTGTCTACTGTGTGTTCTGTGCAAGTGTGTGAATCTTAATAAGA 490
Db      388 CCCCTCTCAGAGGCTCCGTGGGGAAGACATCGGCGCGCCGCTCAAGCTTAACCGCGG 447
Qy      491 GCAACGACATTCGCGGTGAATTTGGGCTGGGGGCTGACACATGCAAAAGATGCCAGGC 550
Db      448 GGAACGCCACATCACCGTCACTGGCGGGCGGCGCTCCACACGCGCAMAGAGCGAGGC 507
Qy      551 ATCTGCTCTGTATAGTCAATGATATGCTTGGCCATCTGGAACCTGTAAGTATCA 610
Db      508 CTCCGGGTTCTGTACTGATCAACGACATGCTCTCGCATCTCTGAGCTCTCAAGTTCA 567
Qy      611 CCAAGGGTGTCTATCATTTGATGATTAATCAACATGATGACGCGGTGAAGAGGCTT 670
Db      568 CAGGGGTGTCTATATGATGATTTGATGATGATGATGATGATGATGATGATGATGATG 627
Qy      671 CTACACACGAGACCGGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 730
Db      628 CTTCACATCAAAACCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 687
Qy      731 AACTGGGGACTTAAGGGAATACCGGGGCTGGCAAGCAAGTATTAATGCTTAATCAACC 790
Db      688 TACTGACATTAATCACTGACGTTGGGGACCCGAGGGAAGCATTAATGCTTAATGTTCC 747
Qy      791 GCTCCGAGACGGGATTTGATGACGAGTCTATGAGGCAATTTTCAAGCCGATCATGTCCA 850
Db      748 CCGAGTGTATGATGATGATGACACACCTTTGTGTGCTGTTTCAATGATCATTAAGAA 807
Qy      851 AGTAATGAGATGATTCAGCTTAGTGCGGTGTTTACAGTGTGCTCAGACTCCATATC 910
Db      808 AGTTATGAGAGTTTATCAACACGACCTGTGTGCTTCCATGCGAGCTGACTCTTGGC 867
Qy      911 TGGGATTCGTTAGTGTGCTTCAATCTATCTATCAAAAGACACGCCAAGTGTGGAAT 970
Db      868 TGGAGACAGTTAGTTAGTTCTTCAACCTGCTGTGAAAGGATCAATGACTGCTCCGTT 927
Qy      1031 CGTTGCCGGTGTGACATATAGACAGCTGTGGCCCTGATATCGAGATCCCTAATGA 1090
Db      988 TGTTCGACGCTGTGCTGCTACGAGACCGCAGTGTGCTTGGAGTTGAACCTGATACAA 1047
Qy      1091 GCTTCCATACATGACTACTTGTGAATATCTTTGGACCAATTTCAAGCTCCACATCAGTCC 1150
Db      1048 GCTGCTTACAAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
Qy      1151 TTCCAAATGACTAACCAACACGAAATGATGATCTGGAGAGATCAAAACAGGACTGTT 1210
Db      1108 AAAAGTGTGAAAACCTGATATCCAAAGGACTTGGAGAACATTAAGAACATGATATT 1167
Qy      1211 TGAAGACTTGAATGCTGCGGACGACCTGGGGTCCAAACGACGAGCAATTCCTGAGA 1270
Db      1168 GAGAACTGTGTCAAAAGATGAAATGATGATGATGATGATGATGATGATGATGATGATG 1227
Qy      1271 CGCCATCCCTGAGAGAGTGGCGATGAGAGCAAGACA 1309
Db      1228 CCTGAAGCTCCAGAGAGAAAGAGAGACATGACAA 1266

```


RESULT 13
US-09-883-720-3
; Sequence 3, Application US/09883720
; Patent No. 6479629
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/883,720
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/282,305
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (29)..(1084)
US-09-883-720-3

Query Match 20.4%; Score 329.2; DB 4; Length 1475;
Best Local Similarity 63.8%; Pred. No. 2.7e-90;
Matches 499; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

QY 528 CACCATGCAAGAAAGCCGAGGATCGCTTCTGTACATGATATCGCTTGCC 587
DB 164 CACCATGCAAGAAAGCCGAGGATCGCTTCTGTACATGATATCGCTTGCC 223
QY 588 ATCTGGAATCTGTAAGTATCAACGAGGCTGTGATGATGATGATGATGATGAT 647
DB 224 ATCTGGAATCTGTAAGTATCAACGAGGCTGTGATGATGATGATGATGATGAT 283
QY 648 GGTGACGCGGTGAGAGGCGCTTCTACACGAGCGGCTGATGATGATGATGATGAT 707
DB 284 GGAGAGTGGGTGAGAGGCGCTTCTACACGAGCGGCTGATGATGATGATGATGAT 343
QY 708 AAGTATGAGAGTCTCCAGGAACTGGGGATCGGAGATGATGATGATGATGATGAT 767
DB 344 AAGTATGAGAGTCTCCAGGAACTGGGGATCGGAGATGATGATGATGATGATGAT 403
QY 768 AAGTATGATCTGTAAGTATCAACGAGGCTGTGATGATGATGATGATGATGAT 827
DB 404 AAGCATTAATGCTGTAAGTATCAACGAGGCTGTGATGATGATGATGATGATGAT 463
QY 828 ATTTTCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 887
DB 464 CTGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 523
QY 888 CAGTGTGCTGAGTCTCCAGGAACTGGGGATCGGAGATGATGATGATGATGATGAT 947
DB 524 CAGTGTGCTGAGTCTCCAGGAACTGGGGATCGGAGATGATGATGATGATGATGAT 583
QY 948 GGACAGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007
DB 584 GGTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 643
QY 1008 GGCGGTGTTTACCACTGTAAGTATCAACGAGCTTGAATGATGATGATGATGAT 1067
DB 644 GGCGGTGTTTACCACTGTAAGTATCAACGAGCTTGAATGATGATGATGATGAT 703
QY 1068 GTGATAGCGAGATCCCTATGATGATGATGATGATGATGATGATGATGATGAT 1127
DB 704 GTTGAAGTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 763
QY 1128 GATTCAAGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1187

DB 764 GATTATCTTCTATATCAACCAAAAAGTTGAAAATCTGATATCAACAAAGACTTG 823
QY 1188 GAGAGATCAAAACGAGCTGTTTGAACCTTGAATGATGATGATGATGATGATGAT 1247
DB 824 GAGAGATCAAAACGAGCTGTTTGAACCTTGAATGATGATGATGATGATGATGAT 883
QY 1248 CAACGAGCGGATCTGAGAGCGCATCCCTGAGAGAGTGGCGATGAGAGAGAG 1307
DB 884 CAATTCATGACAGACCGTCAAGACCTGAAGCTCCAGAGAGAAAGAGAGAGCATGAC 943
QY 1308 GA 1309
DB 944 AA 945

RESULT 14
US-09-389-681-403
; Sequence 403, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 403
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(241)
; OTHER INFORMATION: n = A,T,C or G
US-09-389-681-403

Query Match 14.6%; Score 235.8; DB 4; Length 241;
Best Local Similarity 98.8%; Pred. No. 3.1e-62;
Matches 237; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 777 GCTGTTAACTACCCGCTCCGAGACCGGATGATGATGATGATGATGATGATGAT 836
DB 2 GGTGTTAACTACCCGCTCCGAGACCGGATGATGATGATGATGATGATGATGAT 61
QY 837 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 896
DB 62 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
QY 897 TCAGACTCCCTATCTGGGATGATGATGATGATGATGATGATGATGATGATGAT 956
DB 122 TCAGACTCCCTATCTGGGATGATGATGATGATGATGATGATGATGATGATGAT 181
QY 957 AAGTGTGGAATTTGTCAAGGCTTTAACTGCTATGCTATGCTGAGAGCGGTGT 1016
DB 182 AAGTGTGGAATTTGTCAAGGCTTTAACTGCTATGCTATGCTGAGAGCGGTGT 241

RESULT 15
US-09-620-405B-403
; Sequence 403, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepner, William T.

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 403
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(241)
; OTHER INFORMATION: n = A,T,C or G
US-09-620-405B-403

```

```

Query Match      14.6%; Score 235.8; DB 4; Length 241;
Best Local Similarity 98.8%; Pred. No. 3.1e-62;
Matches 237; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 777 GCTGTTAACTACCCGCTCCGAGAGCGGATGATGACGAGTCTATGAGGCCATTTCAG 836
Db 2 GGTGTAACTAACCCGCTCCGAGAGCGGATGATGACGAGTCTATGAGGCCATTTCAG 61

QY 837 CCGGTCAATGTCCTCAAGTATGAGATGTTCCAGCTAGTGGGTGCTTACAGTGTGGC 896
Db 62 CCGGTCAATGTCCTCAAGTATGAGATGTTCCAGCTAGTGGGTGCTTACAGTGTGGC 121

QY 897 TCAGACTCCCTATCTGGGGATCGGTTAGTGTCTTCAATCTATCTATCAAGAAGACACGCC 956
Db 122 TCAGACTCCCTATCTGGGGATCGGTTAGTGTCTTCAATCTATCTATCAAGAAGACACGCC 181

QY 957 AAGTGTGTGAATTGTGCAAGAGCTTTAAGCTAAGCTGATGCTGGAGGCGGTGGT 1016
Db 182 AAGTGTGTGAATTGTGCAAGAGCTTTAAGCTAAGCTGATGCTGGAGGCGGTGGT 241

```

Search completed: December 5, 2003, 16:41:15
 Job time : 138 secs


```

Db      181  ATGACAAGGCGACCCCAATGAAAGCTCAACGAAATCCGATGACTATTAATTTGCTGCTCA 240
Qy      241  ACTATGCTCTTACGSAATAATGAAATCTATCCGCCCTCAAAAGCCAAATGCTGAGAGA 300
Db      241  ACTATGCTCTTACGSAATAATGAAATCTATCCGCCCTCAAAAGCCAAATGCTGAGAGA 300
Qy      301  TGACCAAGTACACAGCGATGACTATTAATTTCTTGCGCTCCATCCGTCAGATTAACA 360
Db      301  TGACCAAGTACACAGCGATGACTATTAATTTCTTGCGCTCCATCCGTCAGATTAACA 360
Qy      361  TGTGCGAGTACAGCAAGCAGATGAGATTCAAAGTTGGTGGAGAGCTGTCAATTTCCG 420
Db      361  TGTGCGAGTACAGCAAGCAGATGAGATTCAAAGTTGGTGGAGAGCTGTCAATTTCCG 420
Qy      421  ATGGCCCTGTTGAATTTCTGTCAGTTGTTCTACTGGTGTCTGTGAGCAAGCTGTGAAC 480
Db      421  ATGGCCCTGTTGAATTTCTGTCAGTTGTTCTACTGGTGTCTGTGAGCAAGCTGTGAAC 480
Qy      481  TTAATTAAGCAGACGAGCAGATCGCCGTAATTTGGGCTGGGGGCTGCAACATGCAAGA 540
Db      481  TTAATTAAGCAGACGAGCAGATCGCCGTAATTTGGGCTGGGGGCTGCAACATGCAAGA 540
Qy      541  AGTCGAGGCGATCTGGCTTCTGTTACGTCAATGATTCGTTGGCCATCTTGAACTGC 600
Db      541  AGTCGAGGCGATCTGGCTTCTGTTACGTCAATGATTCGTTGGCCATCTTGAACTGC 600
Qy      601  TAAAGTATCACCAAGGGTGTCTGATCATTTGATTTACCAATGATGATGACGGCGTGG 660
Db      601  TAAAGTATCACCAAGGGTGTCTGATCATTTGATTTACCAATGATGATGACGGCGTGG 660
Qy      661  AAGAGGCGCTTATACACCGAGCCGGGTCATGACTGTGTCTTCTTCAATGAATGAGAGT 720
Db      661  AAGAGGCGCTTATACACCGAGCCGGGTCATGACTGTGTCTTCTTCAATGAATGAGAGT 720
Qy      721  ACTTCCAGAGAACTGGGGAGCTTACGGGATACCGGGGCTGGCAAGAAATGATGCTG 780
Db      721  ACTTCCAGAGAACTGGGGAGCTTACGGGATACCGGGGCTGGCAAGAAATGATGCTG 780
Qy      781  TTAATCTACCCGCTCCGAGACGGGATTTGATGACGATGCTTAAAGAGCCATTTCAAGCCG 840
Db      781  TTAATCTACCCGCTCCGAGACGGGATTTGATGACGATGCTTAAAGAGCCATTTCAAGCCG 840
Qy      841  TCATGTCGCAAAATGATGAGATGTTCCAGCTAGTGGCGGTGCTTCAAGTGGCTCAG 900
Db      841  TCATGTCGCAAAATGATGAGATGTTCCAGCTAGTGGCGGTGCTTCAAGTGGCTCAG 900
Qy      901  ACTGCTATCTGGGAGTCGATTAGTGTCTTCAATCTATCTATCAAAAGACCGCAAGT 960
Db      901  ACTGCTATCTGGGAGTCGATTAGTGTCTTCAATCTATCTATCAAAAGACCGCAAGT 960
Qy      961  GTGCGAATTTGTCAAGAGCTTTAACTGCTATGCTGATGCTGGAGGCGGTGTTACA 1020
Db      961  GTGCGAATTTGTCAAGAGCTTTAACTGCTATGCTGATGCTGGAGGCGGTGTTACA 1020
Qy      1021  CCATTCGTAAGTGGCCGGGTGCTGACATATGAGACAGCTGTGGCCCTGATACGAGAGA 1080
Db      1021  CCATTCGTAAGTGGCCGGGTGCTGACATATGAGACAGCTGTGGCCCTGATACGAGAGA 1080
Qy      1081  TCCCTTAATGAGCTTCAATACATGACTTATGATACTTTGAGACGAGATTTCAAGCTCC 1140
Db      1081  TCCCTTAATGAGCTTCAATACATGACTTATGATACTTTGAGACGAGATTTCAAGCTCC 1140
Qy      1141  ACATGAGCTCTTCCAAATGACTTAACGAAACAGAAATGAGTACCTGAGAAATCAAC 1200
Db      1141  ACATGAGCTCTTCCAAATGACTTAACGAAACAGAAATGAGTACCTGAGAAATCAAC 1200
Qy      1201  AGGAGCTGTTGAGAACCTTATGATGCTGCGGACGAGACCTGGGCTCCAAACGAGGCGA 1260
Db      1201  AGGAGCTGTTGAGAACCTTATGATGCTGCGGACGAGACCTGGGCTCCAAACGAGGCGA 1260
Qy      1261  TTCCTGAGAGCGCCATCCCTGAGAGAGTGGCGATGAGAGCAAGACGACCTTGACAAAGC 1320

```

```

Db      1261  TTCCTGAGAGCGCCATCCCTGAGAGAGTGGCGATGAGAGCAAGACGACCTTGACAAAGC 1320
Qy      1321  GCATTCGATCTGCTCCTCTGACAAAGAAATGCTGTGAGAGAGATTCTCGAATTCG 1380
Db      1321  GCATTCGATCTGCTCCTCTGACAAAGAAATGCTGTGAGAGAGATTCTCGAATTCG 1380
Qy      1381  AAGAGAGGAGAGGGGGGCGCAAGAACTTTTCAACTTTCAAAAAAGCCAAAGAGTCA 1440
Db      1381  AAGAGAGGAGAGGGGGGCGCAAGAACTTTTCAACTTTCAAAAAAGCCAAAGAGTCA 1440
Qy      1441  AAACAGAGATGAAAAAGAAAGAACCCAGAGAGAAAGAAAGAAATCAACGAAAGAGAGA 1500
Db      1441  AAACAGAGATGAAAAAGAAAGAACCCAGAGAGAAAGAAAGAAATCAACGAAAGAGAGA 1500
Qy      1501  AAACCAAGAGAGAGAGCCAGAGCCAAAGGGGTCAAGAGAGAGAGCCAAAGTTGGCTGAA 1560
Db      1501  AAACCAAGAGAGAGAGCCAGAGCCAAAGGGGTCAAGAGAGAGAGCCAAAGTTGGCTGAA 1560
Qy      1561  TGGACCTCTTCAGACTCTGGCTTCTGCTGAGTCCCTCAAGTTTCTTTCCCC 1611
Db      1561  TGGACCTCTTCAGACTCTGGCTTCTGCTGAGTCCCTCAAGTTTCTTTCCCC 1611

RESULT 2
US-09-817-538-2
; Sequence 2, Application US/09817538
; Patent No. US20020137162A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zumei
; APPLICANT: Bonifis, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
; TITLE OF INVENTION: Desacylase Isoforms
; FILE REFERENCE: 106101.144
; CURRENT APPLICATION NUMBER: US/09/817,538
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Human
US-09-817-538-2

Query Match      100.0%; Score 1610.6; DB 10; Length 1611;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  ATGTCGGGGTCTCTGCCGCTGGTGTCTGCCACTGGGTATCTGAGAACACA 60
Db      1  ATGTCGGGGTCTCTGCCGCTGGTGTCTGCCACTGGGTATCTGAGAACACA 60
Qy      61  GCCTGAGCGCTCTGTCACTCGGGGTAGACCAAGCGGGAGGCGAGCAAGATGGCGAGA 120
Db      61  GCCTGAGCGCTCTGTCACTCGGGGTAGACCAAGCGGGAGGCGAGCAAGATGGCGAGA 120
Qy      121  CGCAGGCGACCCGAGAGAAATGCTGTTACTATCAAGACGGGAGTGTGAAATTAATTAAT 180
Db      121  CGCAGGCGACCCGAGAGAAATGCTGTTACTATCAAGACGGGAGTGTGAAATTAATTAAT 180
Qy      181  ATGAGCAAGGCGACCAATGAAGCTTCACCGAATTCGCAATGACTCAATTAATTTGCTGCA 240
Db      181  ATGAGCAAGGCGACCAATGAAGCTTCACCGAATTCGCAATGACTCAATTAATTTGCTGCA 240
Qy      241  ACTATGCTCTTACGSAATAATGAAATCTATGCGCTCAAAAGCCAAATGCTGAGAGAGA 300
Db      241  ACTATGCTCTTACGSAATAATGAAATCTATGCGCTCAAAAGCCAAATGCTGAGAGAGA 300
Qy      301  TGACCAAGTACACAGCGATGACTATTAATTTCTTGCGCTCCATCCGTCAGATTAACA 360
Db      301  TGACCAAGTACACAGCGATGACTATTAATTTCTTGCGCTCCATCCGTCAGATTAACA 360

```

```
QY 361 TGTGGAGTACAGACAGATGACAGATTCAACCTTGATGAGACTGTCTCAATTTG 420
DB 361 TGTGGAGTACAGACAGATGACAGATTCAACCTTGATGAGACTGTCTCAATTTG 420
QY 421 ATGGCTGTTTGAAGTCTGTCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 480
DB 421 ATGGCTGTTTGAAGTCTGTCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 480
QY 481 TTAATAGAGAGAGAGAGATGCGCGTGAATTTGGCTGGGGCTGGCACTAGCAAGA 540
DB 481 TTAATAGAGAGAGAGAGATGCGCGTGAATTTGGCTGGGGCTGGCACTAGCAAGA 540
QY 541 AGTCCGAGGACATCTGCTTCTGTATGATGATGATGATGATGATGATGATGATG 600
DB 541 AGTCCGAGGACATCTGCTTCTGTATGATGATGATGATGATGATGATGATGATG 600
QY 601 TAAAGTATACAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 660
DB 601 TAAAGTATACAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 660
QY 661 AAGAGGCTTTTACACACAGAGAGAGTGTGTATGATGATGATGATGATGATG 720
DB 661 AAGAGGCTTTTACACACAGAGAGAGTGTGTATGATGATGATGATGATGATG 720
QY 721 ACTTCCAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 780
DB 721 ACTTCCAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 780
QY 781 TTAATAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 840
DB 781 TTAATAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 840
QY 841 TGAATGCAAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 900
DB 841 TGAATGCAAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 900
QY 901 ACTCCCTATCTGGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 960
DB 901 ACTCCCTATCTGGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 960
QY 961 GTGTGGAATTTTGAAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 1020
DB 961 GTGTGGAATTTTGAAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 1020
QY 1021 CCATTGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
DB 1021 CCATTGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1081 TCCCTAATAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 1140
DB 1081 TCCCTAATAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATG 1140
QY 1141 ACATAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 ACATAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 AGGAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 AGGAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 TTTCCAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1320
DB 1261 TTTCCAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 GCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 AAGAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1440
DB 1381 AAGAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 1440
```

```
QY 1441 AAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1441 AAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 AAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 1501 AAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 TGAAGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1611
DB 1561 TGAAGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1611

RESULT 3
US-09-563-728A-25
; Sequence 25, Application US/09563728A
; Publication No. US20030078216A1
; GENERAL INFORMATION:
; APPLICANT: MacLeod, Alan R
; APPLICANT: Li, Zoumei
; APPLICANT: Besterman, Jeffrey M
; TITLE OF INVENTION: Inhibition of Histone Deacetylase
; FILE REFERENCE: 106101.229
; CURRENT APPLICATION NUMBER: US/09/563,728A
; PRIOR FILING DATE: 2000-05-03
; PRIOR FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 25
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-563-728A-25

Query Match 99.1%; Score 1596.6; DB 11; Length 1611;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1602; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ANGTCTGGGGTCTCTGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
DB 1 ANGTCTGGGGTCTCTGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
QY 61 GCTGAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 61 GCTGAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 121 CGCAGGAGCAGCGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 121 CGCAGGAGCAGCGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 181 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 ACTATGCTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 ACTATGCTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 TGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 TGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 ATGGCTGTTTGAAGTCTGTCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 480
DB 421 ATGGCTGTTTGAAGTCTGTCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 480
QY 481 TTAATAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 540
DB 481 TTAATAGAGAGAGAGAGAGTGTGTATGATGATGATGATGATGATGATGATGATG 540
```


Db 481 TTAATAGCAGACGAGACATCGCTGTGTAATGGGCTGGGCGCTGACCATGCAAGA 540
Qy 541 AGTCGAGGATCTGGCTTCTGTATAGTGAATGATGATGCTTGGCCATCTTGAACTGC 600
Db 541 AGTCGAGGATCTGGCTTCTGTATAGTGAATGATGATGCTTGGCCATCTTGAACTGC 600
Qy 601 TAAAGTATCACCAGAGGGGTGCTGTAACATTGACATTGATATCACCATGATGAGCGCTGG 660
Db 601 TAAAGTATCACCAGAGGGGTGCTGTAACATTGACATTGATATCACCATGATGAGCGCTGG 660
Qy 661 AAGAGGCTTCTACACCGAGCGGCTCATGACTGTGTCTTTCAATAGTAGAGAGT 720
Db 661 AAGAGGCTTCTACACCGAGCGGCTCATGACTGTGTCTTTCAATAGTAGAGAGT 720
Qy 721 ACTTCCAGGAACTGGGGAACTTACCGGAGTACCGGGGCTGGGAAAGACAAAGTATATCTG 780
Db 721 ACTTCCAGGAACTGGGGAACTTACCGGAGTACCGGGGCTGGGAAAGACAAAGTATATCTG 780
Qy 781 TTAATACCGCTCCGAGACGGGATGATGACGAGTCCATGAGGCGCATTTTCAAGCCGG 840
Db 781 TTAATACCGCTCCGAGACGGGATGATGACGAGTCCATGAGGCGCATTTTCAAGCCGG 840
Qy 841 TCATGTCGCAAGATAGAGATGTTCCAGCCTAGTGCCTGCTTCAAGTGTGCTCAG 900
Db 841 TCATGTCGCAAGATAGAGATGTTCCAGCCTAGTGCCTGCTTCAAGTGTGCTCAG 900
Qy 901 ACTCCCTATCTGGGGATCGTTAGGTTGCTTCAATCTATCTATCAAAAGACCCCAAGT 960
Db 901 ACTCCCTATCTGGGGATCGTTAGGTTGCTTCAATCTATCTATCAAAAGACCCCAAGT 960
Qy 961 GTGAGGATTTGTGAAGAGCTTTAACTGCTATGCTGATGCTGGGAGCGGTGTTACA 1020
Db 961 GTGAGGATTTGTGAAGAGCTTTAACTGCTATGCTGATGCTGGGAGCGGTGTTACA 1020
Qy 1021 CCATTCGTAAGCTTCCCGGTGCTGACATATGAGACAGCTGTGCGCTGATACGAGA 1080
Db 1021 CCATTCGTAAGCTTCCCGGTGCTGACATATGAGACAGCTGTGCGCTGATACGAGA 1080
Qy 1081 TCCTTAATGAGCTTCCATCAATGACTACTTGAATCTTTGGACGAGATTTCAAGCTCC 1140
Db 1081 TCCTTAATGAGCTTCCATCAATGACTACTTGAATCTTTGGACGAGATTTCAAGCTCC 1140
Qy 1141 ACATCAGCTCTTCAATATGACTAAACGAGACGAAATGAGTACTGAGAGATCAAAAC 1200
Db 1141 ACATCAGCTCTTCAATATGACTAAACGAGACGAAATGAGTACTGAGAGATCAAAAC 1200
Qy 1201 AGCAGCTGTTGAGAACTTGAATGCTGCGGACGCACTGGGGTCCAAATGCAAGCGA 1260
Db 1201 AGCAGCTGTTGAGAACTTGAATGCTGCGGACGCACTGGGGTCCAAATGCAAGCGA 1260
Qy 1261 TTCTGAGAGAGCCATCCCTGAGAGAGTGGCATGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 TTCTGAGAGAGCCATCCCTGAGAGAGTGGCATGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 GCATCTGATGCTGCTCTCTGACAAAGCAATTCCTGTGAGAGAGAGATTTCTCCATCTG 1380
Db 1321 GCATCTGATGCTGCTCTCTGACAAAGCAATTCCTGTGAGAGAGAGATTTCTCCATCTG 1380
Qy 1381 AAGTCA 1440
Db 1381 AAGTCA 1440
Qy 1441 AAAAGTCA 1500
Db 1441 AAAAGTCA 1500
Qy 1501 AAAAGTCA 1560
Db 1501 AAAAGTCA 1560
Qy 1561 TGAAGCTCTCCAGCTGCTGCTTCTGCTGAGTCCCTCAAGTTTCTTTCCCG 1611
Db 1561 TGAAGCTCTCCAGCTGCTGCTTCTGCTGAGTCCCTCAAGTTTCTTTCCCG 1611

RESULT 4
US-10-133-013-216
; Sequence 216, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
; APPLICANT: Astromoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocke, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
; FILE REFERENCE: PA-0049 US
; CURRENT APPLICATION NUMBER: US/10/133,013
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287,067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PERL Program
; SEQ ID NO 216
; LENGTH: 2098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030166903A1 1513913CB1
US-10-133-013-216

Query Match 93.1%; Score 1499.6; DB 12; Length 2098;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 97 GGGAGGAGACAGATAGGCGCAGACGAGGCGACCCGAGAGAAAGTCTGTTACTACTAG 156
Db 43 GGGAGGAGACAGATAGGCGCAGACGAGGCGACCCGAGAGAAAGTCTGTTACTACTAG 102
Qy 157 ACCGGAGTGTGGAATTAATTAATTAATGAACAGGCCACCCATGAAGCTCAGCGAATCC 216
Db 103 ACCGGAGTGTGGAATTAATTAATTAATGAACAGGCCACCCATGAAGCTCAGCGAATCC 162
Qy 217 GCATGACTCTAATTTGCTGCTCACTATGATGCTCTTACCGAATAAGAAATCTATCGCC 276
Db 163 GCATGACTCTAATTTGCTGCTCACTATGATGCTCTTACCGAATAAGAAATCTATCGCC 222
Qy 277 CTCACAAGCAATGCTGAGAGATGACCAAGTACCAAGAGATGATGATCAATTAATTC 336
Db 223 CTCACAAGCAATGCTGAGAGATGACCAAGTACCAAGAGATGATGATCAATTAATTC 282
Qy 337 TGGGCTTCATCCGTCCAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 396
Db 283 TGGGCTTCATCCGTCCAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 342
Qy 397 TTGGTGAAGACTGCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 456
Db 343 TTGGTGAAGACTGCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 402
Qy 457 GTTCTGAGCAATGCTGGAATCTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 516
Db 403 GTTCTGAGCAATGCTGGAATCTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 462
Qy 517 CTGGGGGCTGACCAATGCAAGAGTCCAGAGCATCTGCTTCTGTTACGTAATGATA 576
Db 463 CTGGGGGCTGACCAATGCAAGAGTCCAGAGCATCTGCTTCTGTTACGTAATGATA 522
Qy 577 TCGTCTTGGCATCTGGAATCTCTTAAGTATACAGAGAGGCTGCTGATGATGATGATG 636
Db 523 TCGTCTTGGCATCTGGAATCTCTTAAGTATACAGAGAGGCTGCTGATGATGATGATG 582
Qy 637 ATATTGACCAATGATGAG 696
Db 583 ATATTGACCAATGATGAG 642
Qy 697 TGTCTTTCAATAGTATGAGAGTACTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
Db 642 TGTCTTTCAATAGTATGAGAGTACTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700

Db 643 TGTCTTTTCAATAGTAGAGAGATCTTCCAGGAACTGAGGAACTTACGGGATATCGGG 702
Qy 757 CTGGCAAGAGCAAGATATATGCTGTTAACTACCCGCTCCGAGCGGGATTTGATGACGAT 816
Db 703 CTGGCAAGAGCAAGATATATGCTGTTAACTACCCGCTCCGAGCGGGATTTGATGACGAT 762
Qy 817 CCTATGAGCCATTTTCAAGCCGGCTATGCTCAAAATGAGATGTTTCAAGCTTCAAGT 876
Db 763 CCTATGAGCCATTTTCAAGCCGGCTATGCTCAAAATGAGATGTTTCAAGCTTCAAGT 822
Qy 877 CGGTGCTTTTCAAGTGTGCTCAAGTCTTATCTGAGGATGCTTGAAGTGTCTTCAATC 936
Db 823 CGGTGCTTTTCAAGTGTGCTCAAGTCTTATCTGAGGATGCTTGAAGTGTCTTCAATC 882
Qy 937 TATCTATCAAGAGCAAGCCAGTGTGTGAATTTTCAAGAGCTTAACTGCTCATG 996
Db 883 TATCTATCAAGAGCAAGCCAGTGTGTGAATTTTCAAGAGCTTAACTGCTCATG 942
Qy 997 TGTATGCTGGAGAGCGGTGTTTACACGATTCGTAACTGTCCTGAGTCTGAGCATATGAG 1056
Db 943 TGTATGCTGGAGAGCGGTGTTTACACGATTCGTAACTGTCCTGAGTCTGAGCATATGAG 1002
Qy 1057 CAGCTGTGCTTGTGATGCGAGATCCCTATATGAGCTTCAATGATGATCTTGAAT 1116
Db 1003 CAGCTGTGCTTGTGATGCGAGATCCCTATATGAGCTTCAATGATGATCTTGAAT 1062
Qy 1117 ACTTTGAGCAAGATTTCAAGTCTCAATGATGCTTCAATGATGATGATGATGATGATG 1176
Db 1063 ACTTTGAGCAAGATTTCAAGTCTCAATGATGCTTCAATGATGATGATGATGATGATG 1122
Qy 1177 ATGAGTACCTGAGAGATCAAAACAGCGACTGTTTGAAGACCTTGAAGTCTGCTGAG 1236
Db 1123 ATGAGTACCTGAGAGATCAAAACAGCGACTGTTTGAAGACCTTGAAGTCTGCTGAG 1182
Qy 1237 CACCTGGGGTCCAAACGAGCGGATTCCTGAGAGCGGATTCCTGAGAGAGTGGCGATG 1296
Db 1183 CACCTGGGGTCCAAACGAGCGGATTCCTGAGAGCGGATTCCTGAGAGAGTGGCGATG 1242
Qy 1297 AGAGCAAGAGCAAGCTGAGAGCGGATTCCTGAGAGCGGATTCCTGAGAGAGTGGCG 1356
Db 1243 AGAGCAAGAGCAAGCTGAGAGCGGATTCCTGAGAGCGGATTCCTGAGAGAGTGGCG 1302
Qy 1357 GTGAGAGAGGATTCCTGAGAGCGGATTCCTGAGAGCGGATTCCTGAGAGAGTGGCG 1416
Db 1303 GTGAGAGAGGATTCCTGAGAGCGGATTCCTGAGAGCGGATTCCTGAGAGAGTGGCG 1362
Qy 1417 ACTTCAAAAAAGCAAGAGATCAAAACAGAGATGAAAAAGAAAAAGCCAGAGAGAG 1476
Db 1363 ACTTCAAAAAAGCAAGAGATCAAAACAGAGATGAAAAAGAAAAAGCCAGAGAGAG 1422
Qy 1477 AGAAAGAGATCAAG 1536
Db 1423 AGAAAGAGATCAAG 1482
Qy 1537 AG 1596
Db 1483 AG 1542
Qy 1597 CACGTTTCTTTCCC 1610
Db 1543 CACGTTTCTTTCCC 1556

RESULT 5

US-09-925-300-623
; Sequence 623, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 623
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-623

Query Match 93.1%; Score 1499.6; DB 10; Length 2163;

Best Local Similarity 99.4%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 97 GGGAGGCGAGCAAGATGGCGAGAGCGAGCGAGCGGAGAGAGCTGTACTACTAG 156
Db 71 GGGAGGCGAGCAAGATGGCGAGAGCGAGCGAGCGGAGAGAGCTGTACTACTAG 130
Qy 157 ACGGGGATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 216
Db 131 ACGGGGATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 190
Qy 217 GCATGATCATTAATTTGCTGCTCAATGATGATGATGATGATGATGATGATGATG 276
Db 191 GCATGATCATTAATTTGCTGCTCAATGATGATGATGATGATGATGATGATGATG 250
Qy 277 CTCAAGAGCGAGATGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGATG 336
Db 251 CTCAAGAGCGAGATGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGATG 310
Qy 337 TGGGCTCCATCCGTCAGATTAATGATGATGATGATGATGATGATGATGATGATG 396
Db 311 TGGGCTCCATCCGTCAGATTAATGATGATGATGATGATGATGATGATGATGATG 370
Qy 397 TTGGTGGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 456
Db 371 TTGGTGGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 430
Qy 457 GTTCTGTGCAAGTGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 516
Db 431 GTTCTGTGCAAGTGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 490
Qy 517 CTGGGGGCTGCAACATGCAAGAGAGTCCGAGGATCTGGCTTCTGTTACGTCATGATA 576
Db 491 CTGGGGGCTGCAACATGCAAGAGAGTCCGAGGATCTGGCTTCTGTTACGTCATGATA 550
Qy 577 TGTCTTTGCGCATCTCGAATCTGTAAGTATCAACAGAGGCTGTATCATTTGACATTG 636
Db 551 TGTCTTTGCGCATCTCGAATCTGTAAGTATCAACAGAGGCTGTATCATTTGACATTG 610
Qy 637 ATTTTCAATGATGAGCGCGGTGGAAGAGAGCTTTCAACACCGGACCGGCTCATAG 696
Db 611 ATTTTCAATGATGAGCGCGGTGGAAGAGAGCTTTCAACACCGGACCGGCTCATAG 670
Qy 697 TGTCTTTTCAATGATGAGAGTACTTCCAGGAAGTGGGAGCTTACCGGATATCGGG 756
Db 671 TGTCTTTTCAATGATGAGAGTACTTCCAGGAAGTGGGAGCTTACCGGATATCGGG 730
Qy 757 CTGGCAAGAGCAAGATTAATGCTGTTAACTACCCGCTCCGAGCGGGATTTGATGACGAGT 816
Db 731 CTGGCAAGAGCAAGATTAATGCTGTTAACTACCCGCTCCGAGCGGGATTTGATGACGAGT 790
Qy 817 CCTATGAGCCATTTTCAAGCCGGTATGCTCAAAATGAGATGTTTCAAGCTTCAAGT 876
Db 791 CCTATGAGCCATTTTCAAGCCGGTATGCTCAAAATGAGATGTTTCAAGCTTCAAGT 850

877 CGGTGATCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTGATGCTTCAATC 936
851 CGGTGATCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTGATGCTTCAATC 910
937 TATCTATCAAGACACGCGCAAGTGTGATTTGTCAAGAGCTTTAACTGCTTATGC 996
911 TAACTATCAAGACACGCGCAAGTGTGATTTGTCAAGAGCTTTAACTGCTTATGC 970
997 TGAATGCTGGGAGGGGGTGTGTACACATTCGTAAAGTTGCGGCTGGGATATGAGA 1056
971 TGAATGCTGGGAGGGGGTGTGTACACATTCGTAAAGTTGCGGCTGGGATATGAGA 1030
1057 CAGCTGTGCGCTCGATACGAGATCCCTATGAGCTTCCATACATGACTACTTGAAT 1116
1031 CAGCTGTGCGCTCGATACGAGATCCCTATGAGCTTCCATACATGACTACTTGAAT 1090
1117 ACTTTGACCAAGATTTCAAGTCCACATCACTGCTTCCATATGACTTACCAAGACGA 1176
1091 ACTTTGACCAAGATTTCAAGTCCACATCACTGCTTCCATATGACTTACCAAGACGA 1150
1177 ATGAGTACCTGGAGAGATCAACACAGGAGCTTTGAGAACCTTAGATGCTGCGGACG 1236
1151 ATGAGTACCTGGAGAGATCAACACAGGAGCTTTGAGAACCTTAGATGCTGCGGACG 1210
1237 CACCTGGGGTCCAAACGAGGCGATCTCTGAGAGCGCCATCCCTGAGAGAGTGGCATG 1296
1211 CACCTGGGGTCCAAACGAGGCGATCTCTGAGAGCGCCATCCCTGAGAGAGTGGCATG 1270
1297 AGACGAAAGACGACCTTGACAAAGCGCATCTGATCTGCTCTGACAAAGAAATTCCT 1356
1271 AGACGAAAGACGACCTTGACAAAGCGCATCTGATCTGCTCTGACAAAGAAATTCCT 1330
1357 GTGAGGAAGATTTCTCCGATTTCTGAAGAGAGAGAGGGGGCGGAAAGACTTTCCA 1416
1331 GTGAGGAAGATTTCTCCGATTTCTGAAGAGAGAGAGGGGGCGGAAAGACTTTCCA 1390
1417 ACTTCAAAAAAGCCAAAGAGATCAAAACAGAGATGAAAAAGAAAGAACCCAGAGAGA 1476
1391 ACTTCAAAAAAGCCAAAGAGATCAAAACAGAGATGAAAAAGAAAGAACCCAGAGAGA 1450
1451 AGAAGAAGATCAAGAGAGAGAGAAAAACCAAGAGAGAGAACCAAGAGAGAGAGTCA 1510
1537 AGGAGAGAGCCAAAGTGGCTGATGAGCTCTCAGCTGCTGCTTCTGCTGAGTCCCT 1596
1511 AGGAGAGAGTCAAGTGGCTGATGAGCTCTCAGCTGCTGCTTCTGCTGAGTCCCT 1570
1597 CACGTTTCTTCCC 1610
1571 CACGTTTCTTCCC 1584

RESULT 6
US-09-754-167-3
Sequence 3, Application US/09754167
Patent No. US20010019328A1
GENERAL INFORMATION:
APPLICANT: Brect P. Monia
APPLICANT: Jacqueline Wyatc
TITLE OF INVENTION: ANTISENSE MODULATION OF HISTONE DEACETYLASE 1 EXPRESSION
FILE REFERENCE: RUS-0140
CURRENT APPLICATION NUMBER: US/09/754.167
CURRENT FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 87
SEQ ID NO 3
LENGTH: 2091
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (64) ... (1512)
US-09-754-167-3

Query Match 93.0%; Score 1498; DB 9; Length 2091;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

97 GGGAGCGAGCAAGATGCGCGACAGCGAGGCAAGCGGAGAAAGTGTACTACTACG 156
50 GGGAGCGAGCAAGATGCGCGACAGCGAGGCAAGCGGAGAAAGTGTACTACTACG 109
157 ACGGGAGTGTGAATTTACTATTATGACAAGAGCCACCAATGAAGCCACCGAATCC 216
110 ACGGGAGTGTGAATTTACTATTATGACAAGAGCCACCAATGAAGCCACCGAATCC 169
217 GCATGACTCATATTTGCTGCTCACTATGCTCTACCGAAAAATGAAATCTATCGCC 276
170 GCATGACTCATATTTGCTGCTCACTATGCTCTACCGAAAAATGAAATCTATCGCC 229
277 CTCAAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGCATGACTTAAATTTCT 336
230 CTCAAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGCATGACTTAAATTTCT 289
337 TGGGCTCCATCCGTCAGATATGATGCGGAGTACCAAGAGATGACAGATTTCAAG 396
290 TGGGCTCCATCCGTCAGATATGATGCGGAGTACCAAGAGATGACAGATTTCAAG 349
397 TTGGTGAAGACTGCTCAGATTTGATGAGCTGTTGAGTTCTGTCAAGTTGTCTACTG 456
350 TTGGTGAAGACTGCTCAGATTTGATGAGCTGTTGAGTTCTGTCAAGTTGTCTACTG 409
457 GTTCTGTGGAAGTGTCTGTAATTTATATAGAGAGAGAGCATGCGCGTAATTTGG 516
410 GTTCTGTGGAAGTGTCTGTAATTTATATAGAGAGAGAGCATGCGCGTAATTTGG 469
517 CTGGGGGCTTGACATGCAAGAGTCCGAGCATCTGCTTCTGTACTAGTAAATATA 576
470 CTGGGGGCTTGACATGCAAGAGTCCGAGCATCTGCTTCTGTACTAGTAAATATA 529
577 TCGCTTGGCCATCTCGAAGCTCTAAAGTATACCAAGAGAGTGTGTACTTGAATTG 636
530 TCGCTTGGCCATCTCGAAGCTCTAAAGTATACCAAGAGAGTGTGTACTTGAATTG 589
637 ATATTCCATGATGAGCGGCTGAGAGAGGCTTTTACACACAGAGCCGGGTATGACTG 696
590 ATATTCCATGATGAGCGGCTGAGAGAGGCTTTTACACACAGAGCCGGGTATGACTG 649
697 TGTCTTTTCAATGATGAGAGTATCTTCCAGAGACTGGGGACTTACCGGATATCGGG 756
650 TGTCTTTTCAATGATGAGAGTATCTTCCAGAGACTGGGGACTTACCGGATATCGGG 709
757 CTGGCAAAAGCAAGTATGCTGTTAACTACCGGCTCCGAGAGCGGATATGATGACAGT 816
710 CTGGCAAAAGCAAGTATGCTGTTAACTACCGGCTCCGAGAGCGGATATGATGACAGT 769
817 CTTATGAGGCAATTTTCAAGCCGCTCATGTCACAAAGTATGAGATGTCACGCTATG 876
770 CTTATGAGGCAATTTTCAAGCCGCTCATGTCACAAAGTATGAGATGTCACGCTATG 829
877 CGGTGCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTATGCTTCAATC 936
830 CGGTGCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTATGCTTCAATC 889
937 TATCTATCAAGACACGCGCAAGTGTGATTTGTCAAGAGCTTTAACTGCTTATGC 996
890 TATCTATCAAGACACGCGCAAGTGTGATTTGTCAAGAGCTTTAACTGCTTATGC 949
997 TGAATGCTGGGAGGGGGTGTGTACCAATTCGTAAAGTTGCGGCTGGGATATGAGA 1056
950 TGAATGCTGGGAGGGGGTGTGTACCAATTCGTAAAGTTGCGGCTGGGATATGAGA 1009
1057 CAGCTGTGCGCTCGATACGAGATCCCTATGAGCTTCCATACATGACTACTTGAAT 1116
1010 CAGCTGTGCGCTCGATACGAGATCCCTATGAGCTTCCATACATGACTACTTGAAT 1069

QY 1117 ACTTGGACCAATTTCAGCTCCATCATGCTCTTCCATATATGATTAACAGACGCA 1176
DB 1070 ACTTGGACCAATTTCAGCTCCATCATGCTCTTCCATATATGATTAACAGACGCA 1129
QY 1177 ATAGATCTCTGGAAGATCAACAGGACTGTTTGAACCTTGAATCTGCGGACG 1236
DB 1130 ATAGATCTCTGGAAGATCAACAGGACTGTTTGAACCTTGAATCTGCGGACG 1189
QY 1237 CACTGGGGTCAAAAGCAGGAGATCTGAGAGCGCATCCCTGAGGAGAGTGGGATG 1296
DB 1190 CACTGGGGTCAAAAGCAGGAGATCTGAGAGCGCATCCCTGAGGAGAGTGGGATG 1249
QY 1297 AGGACGAG 1356
DB 1250 AGGACGAG 1309
QY 1357 GTGAG 1416
DB 1310 GTGAG 1369
QY 1417 ACTTCAAAAAAGCAG 1476
DB 1370 ACTTCAAAAAAGCAG 1429
QY 1477 AGAAGAGATCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536
DB 1430 AGAAGAGATCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489
QY 1537 AG 1596
DB 1490 AG 1549
QY 1597 CAGTTCTTTCC 1610
DB 1550 CAGTTCTTTCC 1563

RESULT 7

US-09-563-728A-27
Sequence 27, Application US/09563728A
GENERAL INFORMATION:
APPLICANT: Macleod, Alan R
APPLICANT: Li, Zoumei
TITLE OF INVENTION: Inhibition of Histone Deacetylase
FILE REFERENCES: 106101.229
CURRENT APPLICATION NUMBER: US/09/563, 728A
PRIOR FILING DATE: 2000-05-03
PRIOR FILING DATE: 1998-05-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 27
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
US-09-563-728A-27

Query Match 50.0%; Score 805.2; DB 11; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3, 1e-237;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 126 GGCACCCGAGGAAAGTCTTACTACTACGACGGGAGTTGGAAATTACTATATGA 185
DB 223 GGCACCCGAGGAAAGTCTTACTACTACTACGACGGGAGTTGGAAATTACTATATGA 282
QY 186 CAGGACCAACCAATGAAGCTCAACGAGATGATGATTAATTGCTGCATCTAT 245
DB 283 CAGGACCAACCAATGAAGCTCAACGAGATGATGATTAATTGCTGCATCTAT 342
QY 246 GGTCTACCGAAATGAAATCTATCGCTCACAAAGCAATGCTGAGAGATGACC 305
DB 1423 TCTATTCGAGATCAGACAAAGGAGATGCTTGTGATGAAGATTTCTCAGATTCTGAGAT 1482

DB 343 GGTCTATACAGAAATGAAATATATAGCCCCCATTAAGCACTGCCCCAGAAATAGCA 402
QY 306 AGTATACAGAGAGATCAATTAATCTTGGGCTCCATCCGTCCAGATTAACATGTG 365
DB 403 AATATACAGAGATCAATTAATCTTGGGCTCCATCCGTCCAGATTAACATGTG 462
QY 366 GAGTACAGAGAGATCAATTAATCTTGGGCTCCATCCGTCCAGATTAACATGTG 425
DB 463 GAGTATAGTACAGATCAATTAATCTTGGGCTCCATCCGTCCAGATTAACATGTG 522
QY 426 CTGTTGAGTCTGCTCAATGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
DB 523 CTGTTGAGTCTGCTCAATGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
QY 486 AAGCAG 545
DB 583 CAGACAG 642
QY 546 GAGGATCTGCTGCTTCTGATCAATGATATCTGCTGAGAGAGAGAGAGAGAGAGAG 605
DB 643 GAGGATCTGAGATCTGCTTCTGATCAATGATATCTGCTGAGAGAGAGAGAGAGAG 702
QY 606 TATCAG 665
DB 703 TATCAG 762
QY 666 GCTTCTTACAG 725
DB 763 GCTTCTTACAG 822
QY 726 CAG 785
DB 823 CAG 882
QY 786 TACCCGCTCCAG 845
DB 883 TACCCGCTCCAG 942
QY 846 TCCAG 905
DB 943 TCCAG 1002
QY 906 CTATCTGAG 965
DB 1003 CTATCTGAG 1062
QY 966 GAATTTGACAG 1025
DB 1063 GAATTTGACAG 1122
QY 1026 CTTAG 1085
DB 1123 CTTAG 1182
QY 1086 AATGAGCTTCAATTAATGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAAT 1145
DB 1183 AATGAGCTTCAATTAATGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAAT 1242
QY 1146 AGTCTTCAATTAATGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1205
DB 1243 AGTCTTCAATTAATGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1302
QY 1206 CTGTTGAG 1265
DB 1303 CTGTTGAG 1362
QY 1266 GAG 1325
DB 1363 GAG 1422
QY 1326 TCGATCTGCTCTGACAAAGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1385
DB 1423 TCGATCTGCTCTGACAAAGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1482

Query 1386 GAGGAGAGGGGGCCGCAAGACTTCCAACTTCAAAAAAGCAAGAGTCAAAACA 1445
Db 1483 GAAGAGAGAGAGGTCCAGAGAAATGTGGCTGATCATTAAGAAAGACAAAGAAAGCTAAGA 1542
Qy 1446 GAGGATGAAAAAGAGAAAGACCAGAGAGAGAAAGAAAGAAATCAAGAGAGAGAAACC 1505
Db 1543 ATTGAAGAGATTAAGAAAGAAAGAGAGACAAAAACAGACGTTAAAGAAAGATTAAGA 1602
Qy 1506 AAGAGAGAGAGCCAGAAAGCCAAAGGGGTCAAGAGAGAGGCCAACT 1551
Db 1603 TCCAAAGGACAAAGTGTGAAAAAACAAGATACCAAGAGAACCAAT 1648

RESULT 8
US-09-373-658-76
; Sequence 76, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Iruela-Arispe, Luisa
; APPLICANT: Haselings, Gregg A.
; APPLICANT: Ruben, Steven W.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trull, Stephen H.
; APPLICANT: Fronwald, James A.
; APPLICANT: Terret, Jonathan A.
; TITLE OR INVENTION: Meth1 and Meth2 polynucleotides and polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373.658
; NUMBER OF SEQUENCE ID NOS: 125
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 76
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Unknown
US-09-373-658-76

Query Match 50.0%; Score 805.2; DB 11; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3.1e-237;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

Qy 126 GCGACCCGAGAGAAAGTGTCTTACTACTACAGACGGGAGTGTGAAAAATTAATTAAGA 185
Db 223 GCGGCAAAAAAAGTGTCTACTACTACAGACGGTATTTGAAAAATTAATTAAGA 282
Qy 186 CAAGGCAACCAATGAAGCTCACCGAATCCGATGATCTCAATTAATTTGCTGCTCAACTAT 245
Db 283 CAGGCTCATCCCATGAGGCTCATAGAAATCCGATGATCCCATTAATTTGCTGCTTAATTA 342
Qy 246 GGTCTTACCGAAAAATGAATCTATGCGCTCAAAAGCCAAATGCTGAGAGATGAC 305
Db 343 GGGTTAATACAAAAATGAATATATAGGCCCAATGAAGCCATGCGAAAGATGACA 402
Qy 306 AAGTACCAACAGATGATTAATTAATTTGCGCTCAATCCGTCGATTAAGATGCG 365
Db 403 AAATATCAAGATGATGATATATCAAAATTTCTACGCTCAATGAAGCCAGATTAAGTCT 462
Qy 426 CTGTTGAGTTCTGCTGATGATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
Db 523 CTCTTGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
Qy 486 AAGGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 545
Db 583 CGAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
Qy 546 GAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605

Db 643 GAAGCATCAGAGATTCCTGTTACGTTAATGATATGTGCTTGCATCTCTGTAATTAAGAAG 702
Qy 606 TATCACAAGAGGAGTCTGATCATTAATGATGATTAATTAATTAATTAATTAATTAATTAAT 665
Db 703 TATCACAAGAGTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 762
Qy 666 GCTTCTACACAG 725
Db 763 GCTTTTATTAACAAGATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 822
Qy 726 CAGAAAG 785
Db 823 CTGAGCAG 882
Qy 786 TACCCGCTCGAAG 845
Db 883 TTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 942
Qy 846 TCCAAAGATTAAGAGATGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 905
Db 943 TCAAGGATTAAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1002
Qy 906 CTATCTGGGAGTGGTTAGGTTGCTTCAATCTATCTATCAAGAGAGAGAGAGAGAGAGAG 965
Db 1003 TTAATCTGATTAAG 1062
Qy 966 GAATTTGTCAAGAGCTTTAATCCGCTGATGATGATGATGATGATGATGATGATGATGAT 1025
Db 1063 GAAGTGTAAACCTTTTAACTTAACATTAATTAATTAATTAATTAATTAATTAATTAAT 1122
Qy 1026 CGTAAGCTGCGGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
Db 1123 CGTAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1182
Qy 1086 AATGAGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1145
Db 1183 AATGAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1242
Qy 1146 AGTCTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1205
Db 1243 AGTCTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1302
Qy 1206 CTGTTTGAAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1265
Db 1303 TTGTTTGAAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
Qy 1266 GAG 1325
Db 1363 GAG 1422
Qy 1326 TCGATCTGCTCTGATCAAG 1385
Db 1423 TCTAATTCAGAGATCAAG 1482
Qy 1386 GAGGAG 1445
Db 1483 GAGGAG 1542
Qy 1446 GAGGATGAAAAAGAGAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1505
Db 1543 ATTGAAGAGATTAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1602
Qy 1506 AAGGAG 1551
Db 1603 TCCAAAGGACAAAGTGTGAAAAAACAAGATACCAAGAGAACCAAT 1648

RESULT 9
US-10-177-293-205
; Sequence 205, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:

APPLICANT: Lillie, James
APPLICANT: Glatc, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Ganavarpur, Manjula
APPLICANT: Kametkar, Shubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Puzsrai, Lajos
APPLICANT: Meric, Punda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 205
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
US-10-177-293-205

Query Match 50.0%; Score 805.2; DB 14; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3.1e-337;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

126 GGACACCGAGGAAAGTCTGTTACTACTACGACGGGAGATGTTGAAATTACTATTATGGA 185
223 GGGGGGAAAAAATCTGCTCTACTGACGCGATATTTGGAATTATTTATGGA 282
186 CAAGGCCACCCAAATGAAGCTCACCAGATCCGATATCTATATTTGCTGCTCAACTAT 245
283 CAGGGTCATCCATGAAGCTCATAGATCCGATATCCATTAATTTGCTGTTAAATTAT 342
246 GGTCTTACCGAAATTTGAAATCTATCGCCCTCAAAAGCCAACTGTGAGGAGAAC 305
343 GGTATTACGAAAAATGAAATATATAGGCCCATTAAGCCACTGCCGAGAAATGACA 402
306 AAGTACCAAGCAGCTACTATTAATTTCTGCGCTCCATCCGTCAGATACATGTG 365
403 AATATACAGATATATATATTAATTTCTACGCTCAATAGACCAATATACATGTCT 462
366 GAGTACAGCAAGAGATGACAGATTCACGTTGGTGAAGACTGTCAGTATTCAGTGGC 425
463 GAGTATAGTATAGATGATATATATTTAATGTTGAGAAATGTCACAGGCTTGAATGA 522
426 CTGTTGAGTCTGCTAGTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
523 CTCTTGAAGTTTGTGACGCTCTCAACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
486 AAGCAGCAGACGAGATCGCGTGAATTGGGCTGGGGCTGTGACCATGCAAGAAAGTCC 545

583 CGACAAACAGCTGATATGCTGTTAATTTGGCTGGAGCATTTACATCATCTTAAGAAATAC 642
546 GAGGATCTGCTGCTGTTTACGTCATGATATGTTGGCATCTGGAACCTGCTAAAG 605
643 GAGGATCAGATTTCTGTTACCTTAATGATATTTGGCTTGCATCTTGAATTAACAAAG 702
606 TATCAACCAAGGGTGTCTGATATGACATTTGATTTACCATGATGACGGCTGGAAG 665
703 TATCATCAGAGAGCTTATATATTTATATGATATTTCAATCATGATGATGTTGAAGAA 762
666 GCTTCTACACCAAGGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
763 GCTTTTATACAAAGATGATATATATGAGTATATTTCAATTAATTTGGGAAATCTTT 822
726 CCAGGAATCGGGGACCTTACGAGATACGGGGCTGGCAAGAAAGATATATGCTGTTAC 785
823 CTTGGCAGAGACCTTGAAGGATATTTGCTGCTGGAAGAAAGCAATATATGCTGCAAT 882
786 TACCGCTCCGAGAGGAGATGATGACAGTCTATGAGGCAATTTCAAGCGGATG 845
883 TTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 942
846 TCCAAAGTATGAGATGTTCCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
943 TCAAGGTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1002
906 CTATCTGAGGAGATCGGTTAGTGTCTTCAATCTATCATCAAGGACAGCCAAAGTGTG 965
1003 TTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1062
966 GAATTTGTCMAAGCTTTTAACTGCTATGATGATGATGATGATGATGATGATGATGATG 1025
1063 GAAGTTGTAATACTTTTAACTTACATGATGATGATGATGATGATGATGATGATGATG 1122
1026 CGTAACTTGGCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1085
1123 CGTAACTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1182
1086 AATGAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1145
1183 AATGAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1242
1146 AGTCTTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1205
1243 AGTCTTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1302
1206 CTGTTTGAACCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265
1303 TTGTTGAATAATTTGGCATGTTACTCTCAAGCCTGGTGTCCAGATGCAAGCTATTTCCA 1362
1266 GAGACGCCATCCCTGAGAGAGTGGCATGAGACGAAGACGACCTGTCAGAGCCATC 1325
1363 GAGAGTCTGTTCTATGAAACAGTGAATGAAGATGAGAGATCAAGCAAGAAATTT 1422
1326 TCGATCTGCTCTGATCAAAAGATTTGCTGAGAGAGATTTCCGATTTGGAAGAG 1385
1423 TCTATTGACATCAGACAGAGGATGCTTGTGATGAAAGATTTCTCAATTTCTGAGAT 1482
1386 GAGGAGAGGGGGCGCAGAACTTTTCAATTTCAAAAAAGCCAAAGAGATCAAAACA 1445
1483 GAAAGGAGGAGAGGTGGAAGAAATGTGGCTGATCATTAAGAAAGACAAAGAAAGTGA 1542
1446 GAGGATGAAAAAGAAAGAACCCAGAGAGAGAAAGAAATCAAGAGAGAGAAAGAAC 1505
1543 ATTGAAGAAATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1602
1506 AAGGAGAGAGGAG 1551
1603 TCCAAAGCAACAGTGTGTAATAAAGATTAACAAAGAAAGCAAT 1648

RESULT 10
US-09-817-913-4

Sequence 4, Application US/09817913
Patent No. US20020061860A1
GENERAL INFORMATION:
APPLICANT: Li, Zuomei
APPLICANT: Bonifis, Claire
APPLICANT: Besteman, Jeffrey
TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
FILE REFERENCE: 106101.145
CURRENT APPLICATION NUMBER: US/09/817,913
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,157
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1985
TYPE: DNA
ORGANISM: Human
US-09-817-913-4

Query Match 49.4%; Score 795.6; DB 9; Length 1985;
Best Local Similarity 72.4%; Pred. No. 2,9e-234;
Matches 1032; Conservative 0; Mismatches 394; Indels 0; Gaps 0;

126 GGCACCCGAGAGAGTCTGTTACTACTACGACGGGATGTTGGAATTAATTATATGA 185
223 GGGCGCAAAAAAAGTCTGTTACTACTACGACGGTGTATATGGAATTAATTATGA 282
186 CAAGGCCACCAATGAAGCTGACCGAATCCGATGATCTCATTAATTTGCTCAACTAT 245
283 CAGGGTCATCCCATGAGACCTCATAGAAATCCGATGACCCATTAATTTGCTCAACTAT 342
246 GGTCTTACCGAATAAAGTATCTATCGCCCTACAAAGCCATGCTGAGAGATACC 305
343 GGGTTACACAGAAATGAATATATAGCCCATTAAGCCATGCTGAGAGATATCA 402
306 AAGTACACAGCGATGATCATTAATTTCTTGGCTCCATCCGTCAGATTAATGCTG 365
403 AATATACAGTGTATGATATATCAATTTCTACGTCATTAACACGATTAATGCTG 462
366 GAGTACAGCAAGCATGATGACAGATTCACGTTGGTGAAGATGCTGCAATTTGATGC 425
463 GAGTATAGTAAAGCATGATATATTAATTTGAGAAAGATTTGTCAGCGTTGATGGA 522
426 CTGTTTGAATTTCTGATGTTCTACTGTTGTTCTGTTGCAATGCTGTTGAATTAAT 485
523 CTCTTTAGTTTGTGAGCTTCAACTGGCGTTCACTGCTGAGCTGTTGAATTAAT 582
486 AAGCAGAGACGACATGCGCGTGAATTTGGGCTGGGGCCCTGACACATGCAAAAGATCC 545
583 CGACACACAGCTGATATGCTGTTAATTTGGCTGGAGGATTAATCATGCTTAAGAAATAC 642
546 GAGGATCTGGCTTCTGTTACGATATATGCTTCTTGGCCATCTGGAATCTGTAAG 605
643 GAACATCAGGATCTGTTACGTTAATGATATGCTGTCATCTGTAATTAATTAAG 702
606 TATACACAGAGGCTGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 665
703 TATCATAGAGAGCTTATATATGATATTAATTAATTAATTAATTAATTAATTAAT 762
666 GCCTTCAACCAAGGACCGGGTATGATGCTGCTTCAATTAATTAATTAATTAATTAAT 725
763 GCTTTTATTAACAGATCGTGTATGACGGTATATTCATTAATTAATTAATTAATTAAT 822
726 CCAGAACTGGGACCTACGAGATACCGGGCTGCAAAAGCAATTAATTAATTAATTAAT 785
823 CTTGGACAGAGACTTGAGGATATTTGCTGTAAGAAAGCAATTAATTAATTAATTAAT 882
786 TACCCGCTCCAGACCGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 845
883 TTTCCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 942
846 TCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905

943 TCAAGGATGAGATGATATCAACTAGTGTGTATTAAGATGATGATGATGATGATGAT 1002
906 CTATCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
1003 TATCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1062
966 GAATTTGTCAGAGCTTTTAACCTGATGATGATGATGATGATGATGATGATGATGAT 1025
1063 GAAGTTGTAACCTTTTAACCTGATGATGATGATGATGATGATGATGATGATGAT 1122
1026 CGTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085
1123 CGTATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1182
1086 AATGACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1145
1183 AAGGATGCTCATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
1146 AGTCTTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205
1243 AGTCTTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
1206 CTGTTGAGAACTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265
1303 TTGTTGAGAAATTTGGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
1266 GAGAGCCCATCTCCTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1325
1363 GAAGATGCTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1422
1326 TCGATGCTCTCTGACAAACGAATTTGCTGAGAGAGATGATGATGATGATGATGAT 1385
1423 TCTATTCAGATCATGACAAACGATGATGATGATGATGATGATGATGATGATGATGAT 1482
1483 GAAG 1542
1446 GAGATGAGAAAG 1505
1543 ATTGAAGAGATGAT 1602
1506 AAGGAT 1551
1603 TCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1648

RESULT 11
US-09-817-538-4
Sequence 4, Application US/09817538
Patent No. US20020137162A1
GENERAL INFORMATION:
APPLICANT: Li, Zuomei
APPLICANT: Bonifis, Claire
APPLICANT: Besteman, Jeffrey
TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
FILE REFERENCE: 106101.144
CURRENT APPLICATION NUMBER: US/09/817,538
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,157
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1985
TYPE: DNA
ORGANISM: Human
US-09-817-538-4

Query Match 49.4%; Score 795.6; DB 10; Length 1985;
Best Local Similarity 72.4%; Pred. No. 2,9e-234;

FILE REFERENCE: 5718-44,
 CURRENT APPLICATION NUMBER: US/09/883,720
 CURRENT FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: 09/282,305
 PRIOR FILING DATE: 1999-03-31
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 7
 LENGTH: 1943
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (57) .. (1610)
 US-09-883-720-7

Query Match 31.2%; Score 502.6; DB 9; Length 1943;
 Best Local Similarity 65.5%; Pred. No. 6,9e-144;
 Matches 768; Conservative 0; Mismatches 399; Indels 6; Gaps 2;

134 GAGGAAGTCTGTACTACTAGACGGGGATGTTGAATTTATATATGGAACAAGGCA 193
 122 GGGGGCGGTGTACTTCTAGACCCGAGTGGGCACTACTACTAGCGGCGGCA 181
 194 CCCAATGAAGCCTCACCGAATCCGATGACTCATATTTGCTGCTCAATGATGCTCTA 253
 182 TCCGATGAAGCCGACCGCATCCGATGAGCACTGCTGCTGGCGGCTACGGCTCT 241
 254 CCGAAAATGAAATCTATTCGCTCACAAGCCAAATGCTGAGAGATGACCAATGACA 313
 242 CAACGAGATGAGTGTACCGCCCAACCGGCGCCGACCGGACCTCTGCGCTTCA 301
 314 CAGCGATGACTAATTAATTTGGGCTCCATCCGCTCAAGATAATGTCGAGTACAG 373
 302 CGCGGCGATCATCACTTCTGTGCTCCGCTCAGCGGAAACGAGAGACAGAT 361
 374 CAGCAGATGACAGATTCACGTTGGTGAAGCTGCTCCAGTTCGATGAGCTGTTGA 433
 362 CGCGCTGCTCAAGCGCTTCAAGCTCGGCAAGACTCCCGCTTTCAGCGGCTTACAG 421
 434 GTTCTGATGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
 422 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
 494 GAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
 482 ---TGACATGCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
 554 TGGCTCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 539 GGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
 614 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
 599 GAGAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
 674 CACCAAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
 659 CACCAAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
 734 TGGGACCTACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
 719 AGGGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
 794 CCGAGACGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
 779 AGATGATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 838
 854 AATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 839 TATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898
 914 GATGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973

1034 TGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
 1019 TGCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
 1094 TCCATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
 1079 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1138
 1154 CATATGACTACCAAGACCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1213
 1139 TACATGAGAACCAAAATACCAAGACCAAGATGCTGCTGCTGCTGCTGCTGCTGCT 1198
 1214 GAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
 1199 TATCTTTTAAATTTTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258
 1271 CGCATTCCTGAGAGAGTGGGATGAGAGCA 1303
 1259 AGAATACCTGAGCAAGATGAGATCAAGATGA 1291

RESULT 15
 US-09-883-720-5
 ; Sequence 5, Application US/09883720
 ; Patent No. US200202256A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/883,720
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 09/282,305
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 2019
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (140) .. (1459)
 US-09-883-720-5

Query Match 31.1%; Score 501.8; DB 9; Length 2019;
 Best Local Similarity 65.3%; Pred. No. 1.2e-143;
 Matches 753; Conservative 0; Mismatches 397; Indels 3; Gaps 1;

121 CCGAGGCAACCGGAGAAAGTCTGTTACTAGACGGGATGTTGAATTTACTATT 180
 189 CCGAGCGGTGGAAGCCCGCTGCTGCTTACACCGGAGTGGGCACTACTACT 248
 181 ATGACAAAGGCCCAACCAATGAAGCCTCACCGAATCCGATGCTGCTGCTGCTGCT 240
 249 ACAGGCAAGGCCCAACCGATGAAGCCGACCGCATTCGATGACCAAGCGGCTGCTGCT 308
 241 ACTATGCTCTTACCGAATAATGAAATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 309 GCTACGCGCTCTTCAAGCAAGATGAAAGTTCGAGCTTCAACCTGCGGAGACCGGAGCC 368
 301 TGAACAGTACCAAGAGATGACTATTAATTTTGGGCTGCTGCTGCTGCTGCTGCTGCT 360
 369 TCTGCGCTTCAACCGGAGATGAGTACGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 428

